Étude probabiliste de systèmes de particules en interaction : applications à la simulation moléculaire

Raphaël Roux

UNIVERSITÉ PARIS-EST

Thèse présentée pour obtenir le grade de

Docteur de l'Université Paris-Est

Spécialité : Mathématiques

par

Raphaël Roux

Étude probabiliste de systèmes de particules en interaction. Applications à la simulation moléculaire.

Thèse soutenue le 6 décembre 2010 devant le jury composé de :

Anders Szepessy	*Rapporteur*
Denis Talay	*Rapporteur*
Roland Assaraf	*Examinateur*
Nicolas Fournier	*Examinateur*
Clément Mouhot	*Examinateur*
Benjamin Jourdain	*Directeur de thèse*
Tony Lelièvre	*Directeur de thèse*

À Maud,
à Thomas.

Remerciements

Je tiens tout d'abord à remercier Benjamin Jourdain et Tony Lelièvre pour avoir accepté d'encadrer ma thèse et m'avoir proposé ce sujet passionnant. Je les remercie tout particulièrement pour la patience et la disponibilité dont ils ont fait preuve tout au long de ces trois années passées au CERMICS.

Je remercie également Anders Szepessy et Denis Talay qui m'ont fait l'honneur de rapporter cette thèse, ainsi que Roland Assaraf, Nicolas Fournier et Clément Mouhot, qui ont accepté de faire partie du jury.

Le CERMICS a été un cadre propice pour préparer une thèse pendant trois ans, et je tiens à remercier toutes les personnes qui auraient croisé mon chemin pendant cette période. Je pense aux permanents du CERMICS et aux chercheurs de l'université de Marne-La-Vallée que j'ai eu l'occasion de rencontrer : Aurélien Alfonsi, Éric Cancès, Djalil Chafaï, Jean-Philippe Chancelier, Ismaïla Dabo, Michel De Lara, Jean-François Delmas, Alexandre Ern, Bernard Lapeyre, Claude Le Bris, Frédéric Legoll, Miguel Martinez, Régis Monneau, Gabriel Stoltz, Pierre Vandekerkhove, Je pense également aux chercheurs que j'ai rencontré à Rennes et qu'il m'arrive de revoir au détour d'une conférence : Jean Baptiste Bardet, Arnaud Debussche, Florent Malrieu.

Mes remerciements vont également à tous les doctorants et post-doctorants du CERMICS et du LAMA qui ont partagé avec moi cette période de joies et de galères qu'est la thèse : Abdel, Arnaud, Christina, David D., David P., Jérôme, José, Julie, Kimiya, Laurent D., Laurent M., Matthew, Maxence, Mohamed, Nadia, Olivier, Pierre, Rémi, Ronan, Salma, Simone, Stefano, Virginie. Enfin, je remercie également les secrétaires du laboratoire et de l'Université Paris-Est pour leur précieux travail d'organisation : Catherine Baccaert, Sylvie Berte, Sylvie Cach, Martine Ouhanna et Nathalie Quelleu.

Je tiens également à remercier les enseignants qui ont accepté de me confier des cours et travaux dirigés à l'ESIEE, l'École des Ponts et l'ENSTA, et, dans le cadre de mon monitorat, à l'Université Paris-Est Marne-La-Vallée, ainsi que ceux qui auraient assuré ces cours à mes côtés : Bénédicte Blanc, Marco Cannone, Pierre Chantelot, Claudine Degand, Lionel Dutheil, Frédéric Praslon, Nathaël Gozlan, Magadalena Kobylanski, Guillaume Poly, Paul-Marie Samson, Yijun Xiao.

Je tiens enfin à exprimer toute ma gratitude à mes proches, famille et amis, en particulier à mes parents, pour leur soutien.

Et bien sûr un grand merci à Maud sans qui tout ceci n'aurait pas de sens, et au petit Thomas que son papa embrasse tendrement.

Systèmes de particules en interaction probabiliste. Applications à la simulation moléculaire

Résumé : Ce travail présente quelques résultats sur les systèmes de particules en interaction pour l'interprétation probabiliste des équations aux dérivées partielles, avec des applications à des questions de dynamique moléculaire et de chimie quantique. On présente notamment une méthode particulaire permettant d'analyser le processus de la force biaisante adaptative, utilisé en dynamique moléculaire pour le calcul de différences d'énergies libres. On étudie également la sensibilité de dynamiques stochastiques par rapport à un paramètre, en vue du calcul des forces dans l'approximation de Born-Oppenheimer pour rechercher l'état quantique fondamental de molécules. Enfin, on présente un schéma numérique basé sur un système de particules pour résoudre des lois de conservation scalaires, avec un terme de diffusion anormale se traduisant par une dynamique de sauts sur les particules.

Mots-clés : Systèmes de particules en interaction probabiliste, interprétation probabiliste des équations aux dérivées partielles, calculs d'énergies libres, simulation moléculaire, méthodes de Monte Carlo en chimie quantique, processus de Lévy, lois de conservation hyperboliques.

Probabilistic interacting particle system and application to molecular simulation

Abstract : This work presents some results on stochastically interacting particle systems and probabilistic interpretations of partial differential equations with applications to molecular dynamics and quantum chemistry. We present a particle method allowing to analyze the adaptive biasing force process, used in molecular dynamics for the computation of free energy differences. We also study the sensitivity of stochastic dynamics with respect to some parameter, aiming at the computation of forces in the Born-Oppenheimer approximation for determining the fundamental quantum state of molecules. Finally, we present a numerical scheme based on a particle system for the resolution of scalar conservation laws with an anomalous diffusion term, corresponding to a jump dynamics on the particles.

Keywords : Interacting particle systems, probabilistic interpretation of partial differential equations, free energy calculations, molecular dynamics, quantum Monte Carlo methods, Lévy processes, hyperbolic conservation laws.

AMS Classification : 35K55, 35L60, 35Q84, 60G51, 65C30, 65C35

Table des matières

1

Préambule

Cette thèse regroupe des travaux portant sur l'utilisation des systèmes de particules en interaction pour l'interprétation probabiliste des équations aux dérivées partielles.

En introduction, on donne des exemples classiques d'interprétation probabiliste des équations aux dérivées partielles. Les fondements de la physique statistique et de la physique quantique, dont nous utiliserons le formalisme, sont également rappelés.

La partie II traite d'une approximation particulaire pour la méthode de la force biaisante adaptative (ABF en anglais), utilisée en physique statistique pour le calcul d'énergies libres. Ce travail a été publié dans *Modélisation Mathématique et Analyse Numérique*, voir [37].

Dans la partie III, on étudie la sensibilité de processus de diffusion par rapport à un paramètre, en vue du calcul de configurations électroniques en chimie quantique.

La partie IV présente une méthode numérique pour résoudre une équation aux dérivées partielles hyperbolique non linéaire avec une diffusion anormale. Cette méthode est basée sur un système particulaire suivant une dynamique de sauts. Ce travail a été soumis à *Stochastic Processes and their Applications*.

Partie I

Introduction

2

Interprétation probabiliste des équations aux dérivées partielles

Les équations aux dérivées partielles jouent un rôle central en mathématiques appliquées, car elles permettent de représenter toutes sortes de phénomènes, allant de la diffusion de la chaleur à l'évolution des particules quantiques, en passant par le mouvement des fluides et les prix d'options financières. Il est donc crucial d'être capable de résoudre ces équations de manière performante.

2.1 Quelques exemples d'interprétations probabilistes

L'interprétation probabiliste des équations aux dérivées partielles permet de résoudre ces équations en évitant le recours aux méthodes déterministes classiques, qui peuvent être très coûteuses, voire impossibles, lorsque la dimension du problème considéré est grande. Le principe est d'exprimer la solution d'une équation aux dérivées partielles comme un objet probabiliste, typiquement une espérance ou une densité de probabilité, dépendant de la solution d'un problème probabiliste, le plus souvent une équation différentielle stochastique. On dispose ensuite d'outils classiques de théorie des probabilités pour simuler cet objet probabiliste, par exemple les méthodes de Monte Carlo pour calculer les espérances.

Comme référence sur ces différentes interprétations probabilistes, on pourra consulter [40].

2.1.1 Problèmes elliptiques

Prenons un exemple simple d'une telle interprétation : considérons l'équation de la chaleur

$$\begin{cases} \partial_t u_t(x) = & \frac{1}{2}\Delta u_t(x), \quad (x,t) \in \mathbb{R}^d \times (0,\infty) \\ u_0 & \text{donné} \end{cases}, \tag{2.1}$$

où l'on note $\Delta = \sum_i \partial_i^2$ l'opérateur Laplacien. Cette équation aux dérivées partielles est liée au mouvement Brownien par la formule d'Itō. En effet supposons que la condition initiale u_0 soit une mesure de probabilité, et considérons un mouvement Brownien $(W_t)_{t\geq 0}$ dont la condition initiale W_0 suit la loi u_0. Étant donnée une fonction test φ suffisamment régulière, la formule d'Itō appliquée à $\varphi(W_t)$ donne

$$\varphi(W_t) - \varphi(W_0) = \int_0^t \nabla\varphi(W_s)\mathrm{d}W_s + \frac{1}{2}\int_0^t \Delta\varphi(W_s)\mathrm{d}s.$$

Le premier terme du second membre est une intégrale stochastique par rapport à la martingale W_t, et est donc lui aussi une martingale, puisque la fonction φ est régulière. En conséquence, un passage à l'espérance donne

$$\mathbb{E}[\varphi(W_t)] - \mathbb{E}[\varphi(W_0)] = \frac{1}{2}\mathbb{E}\int_0^t \Delta\varphi(W_s)\mathrm{d}s. \tag{2.2}$$

En notant u_t la loi de la variable aléatoire W_t, on peut récrire l'égalité (2.2) comme

$$\int_{\mathbb{R}} \varphi u_t - \int_{\mathbb{R}} \varphi u_0 = \frac{1}{2} \int_0^t \int_{\mathbb{R}} \Delta \varphi u_s \mathrm{d}s,$$

ce qui est une formulation faible de l'équation de la chaleur (2.1). Par conséquent, la solution de l'équation aux dérivées partielles (2.1) peut s'obtenir comme la densité de la loi d'un mouvement Brownien dont la condition initiale suit la loi u_0.

De manière plus générale, si $(X_t)_{t\geq 0}$ est un processus de Markov de générateur infinitésimal $\mathcal{L}$ à valeurs dans un espace d'états $\mathcal{E}$, ayant pour condition initiale une loi u_0, alors la loi u_t de X_t à l'instant t est donnée par la solution de l'équation aux dérivées partielles suivante, appelée *équation de Kolmogorov forward* ou *équation de Fokker-Planck*

$$\begin{cases} \partial_t u_t(x) = & \mathcal{L}^* u_t(x), \quad (x,t) \in \mathcal{E} \times (0,\infty) \\ u_0 & \text{donné} \end{cases}, \tag{2.3}$$

où $\mathcal{L}^*$ est l'adjoint de l'opérateur $\mathcal{L}$. Notamment, les équations aux dérivées partielles paraboliques entrent dans ce cadre, dans le cas où l'opérateur $\mathcal{L}$ s'écrit

$$\mathcal{L} = \sum_{i=1}^n b_i(x)\partial_i + \frac{1}{2} \sum_{i,j=1}^n a_{ij}(x)\partial_i \partial_j.$$

Sous de bonnes hypothèses de régularité et de non explosion sur les fonctions a et b, cet opérateur est le générateur de la solution de l'équation différentielle stochastique

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \tag{2.4}$$

où $\sigma\sigma^T = a$.

On peut également donner une autre interprétation probabiliste à l'équation (2.1). En effet, en appliquant la formule d'Itō au processus $u_0(x + W_t)$, on trouve

$$u_0(x + W_t) - u_0(x) = \int_0^t \nabla u_0(x + W_s)\mathrm{d}W_s + \frac{1}{2} \int_0^t \Delta u_0(x + W_s)\mathrm{d}s,$$

d'où, en passant à l'espérance

$$\mathbb{E}[u_0(x + W_t)] = u_0(x) + \int_0^t \Delta \mathbb{E}[u_0(x + W_s)]\mathrm{d}s,$$

montrant que $\mathbb{E}[u_0(x + W_t)]$ est une solution de l'équation (2.3). Ce résultat se généralise aux équations de la forme

$$\begin{cases} \partial_t u_t(x) = & \mathcal{L} u_t(x), \quad (x,t) \in \mathcal{E} \times (0,\infty) \\ u_0 & \text{donné} \end{cases}, \tag{2.5}$$

où $\mathcal{L}$ est le générateur infinitésimal d'un processus de Markov. L'équation (2.5) est appelée *équation de Kolmogorov rétrograde*. La solution de l'équation (2.5) est donnée par $\mathbb{E}[u_0(X_t^x)]$ où $(X_t^x)_{t\geq 0}$ est un processus de Markov de générateur $\mathcal{L}$ issu de x.

2.1.2 Problèmes paraboliques avec conditions aux bords

Considérons un autre type d'équation, le problème de Dirichlet

$$\begin{cases} \Delta u(x) &= 0, \quad x \in \Omega \\ u(x) &= f(x), \quad x \in \partial\Omega \end{cases}, \tag{2.6}$$

où Ω est un ouvert borné de $\mathbb{R}^n$, et f est une fonction bornée définie sur $\partial\Omega$. Supposons que cette équation admette une solution régulière u. En considérant un mouvement Brownien $(W_t^x)_{t\geq 0}$ issu d'un point x de Ω, et en notant τ le temps d'atteinte de $\partial\Omega$, qui est presque sûrement fini puisque Ω est borné, on peut écrire, par la formule d'Itō,

$$u(W_{t\wedge\tau}^x) = u(x) + \frac{1}{2}\int_0^{t\wedge\tau} \Delta u(W_s^x)\mathrm{d}s + \int_0^{t\wedge\tau} \nabla u(W_s^x)\mathrm{d}W_s^x.$$

Ce qui donne, en passant à l'espérance, puis en prenant la limite t tendant vers l'infini,

$$\mathbb{E}[f(W_\tau^x)] = u(x).$$

Par conséquent, la solution régulière du problème de Dirichlet (2.6) s'exprime comme l'espérance de f contre la loi du mouvement Brownien arrêté au bord de Ω.

Cette interprétation se généralise à un processus de Markov $(X_t)_{t\geq 0}$ de générateur $\mathcal{L}$. En effet, considérons un sous ensemble $\mathcal{D}$ de l'espace d'état $\mathcal{E}$. Si une trajectoire de $(X_t)_{t\geq 0}$ issue d'un point quelconque de $\mathcal{E}$ atteint presque sûrement $\mathcal{D}$ en un temps fini, alors, pour une fonction f bornée, toute solution du problème

$$\begin{cases} \mathcal{L}u(x) &= 0, \quad x \in \mathcal{E} \\ u(x) &= f(x), \quad x \in \mathcal{D} \end{cases},$$

s'exprime comme $\mathbb{E}[f(X_\tau^x)]$ où $(X_t^x)_{t\geq 0}$ est un processus de Markov de générateur $\mathcal{L}$ issu de x, et τ est le temps d'atteinte de $\mathcal{D}$.

2.1.3 La formule de Feynman-Kac

On a aussi une interprétation probabiliste pour les équations incluant un terme linéaire multiplicatif, qui correspond à une dissipation locale :

$$\begin{cases} \partial_t u_t(x) &= \mathcal{L}u_t(x) - V(x)u_t(x), \quad (x,t) \in \mathcal{E}\times(0,\infty) \\ u_0(x) &= f(x), \quad x \in \mathcal{E} \end{cases}. \tag{2.7}$$

Soit $(X_t^x)_{t\geq 0}$ un processus de Markov de générateur $\mathcal{L}$ issu du point x. Si u est une solution régulière de l'équation (2.7), en appliquant la formule d'Itō au processus $(u_{t-s}(X_s^x)e^{-\int_0^s V(X_u^x)\mathrm{d}u})_{0\leq s\leq t}$, on obtient la formule donnant l'interprétation probabiliste de (2.7), connue sous le nom de formule de Feynman-Kac :

$$u_t(x) = \mathbb{E}\left[f(X_t^x)e^{-\int_0^t V(X_s^x)\mathrm{d}s}\right]. \tag{2.8}$$

Par conséquent, toute solution régulière de (2.7) s'écrit nécessairement sous la forme (2.8). Cette interprétation est une écriture de u_t comme la loi de X_t, pondérée par un poids exponentiel effectuant une moyenne de l'énergie du processus sur toute sa trajectoire passée.

Ce type d'interprétation probabiliste est notamment à la base de la méthode de Monte Carlo diffusive en chimie quantique.

2.2 Systèmes de particules en interaction probabiliste

Les systèmes de particules en interaction sont des ensembles de processus, appelés "particules" ou encore "marcheurs" dans la littérature physique, dont le mouvement est régi par deux

phénomènes. D'une part, les particules suivent de manière indépendante une même dynamique Markovienne, mais d'autre part cette dynamique est modifiée par une interaction entre les particules, faisant que chaque particule est influencée par l'ensemble du système.

Les systèmes de particules en interaction peuvent servir aussi bien pour des motifs théoriques que numériques. Dans un cadre théorique, ils permettent d'étudier des équations aux dérivées partielles non linéaires. Du côté numérique, ils permettent par exemple de réduire la variance dans le cadre d'une formule de Feynman-Kac, dont le terme exponentiel tend à avoir une variance très importante, voir par exemple [27, 49, 59, 60].

2.2.1 Interprétation des équations aux dérivées partielles non linéaires

On remarquera que les équations (2.3) et (2.5) sont linéaires par rapport à la fonction inconnue u. Le processus stochastique associé sera donc qualifiée de *linéaire au sens de McKean*, même si son évolution est régie par une équation différentielle stochastique dont les coefficients ne sont pas linéaires par rapport au processus $(X_t)_{t\geq 0}$.

La question qui se pose alors naturellement est de trouver une telle interprétation probabiliste à des équations aux dérivées partielles semblables dont certains coefficients ne sont pas linéaires. Par exemple, considérons l'équation de McKean-Vlasov, introduite par McKean dans [53] :

$$\begin{cases} \partial_t u_t(x) = & -\partial_x(B[x,u_t]u_t(x)) + \frac{1}{2}\Delta\left(\Sigma[x,u_t]\Sigma[x,u_t]^T u_t(x)\right), \quad (x,t)\in \Omega\times(0,\infty) \\ u_0(x) & \text{donné} \end{cases},$$

où on note

$$B[x,u_t] = \int_{\mathbb{R}} b(x,y)\mathrm{d}u_t(y), \text{ et } \Sigma[x,u_t] = \int_{\mathbb{R}} \sigma(x,y)\mathrm{d}u_t(y).$$

En appliquant la formule d'Itō comme précédemment, on voit que u_t peut s'interpréter formellement comme la loi d'un processus $(X_t)_{t\geq 0}$ satisfaisant l'équation

$$\begin{cases} \mathrm{d}X_t &= \displaystyle\int_{\mathbb{R}} b(X_t,y)\mathrm{d}u_t(y)\mathrm{d}t + \int_{\mathbb{R}} \sigma(X_t,y)\mathrm{d}u_t(y)\mathrm{d}W_t \\ u_t &= \mathrm{Loi}(X_t) \end{cases}, \tag{2.9}$$

voir par exemple [54]. Il y a une différence fondamentale entre les équations (2.4) et (2.9) : en effet les coefficients de cette dernière dépendent non seulement de la position X_t de la solution à l'instant t, mais aussi de toute la loi de X_t, ce qui fait qu'il est par exemple impossible de traiter ce genre d'équation par des méthodes trajectorielles. Ce type d'équation différentielle stochastique sera qualifié de *non linéaire au sens de McKean*, le terme "non linéaire" traduisant le fait que l'équation aux dérivées partielles associée est non linéaire.

Comment peut-on montrer des résultats d'existence pour les équations (2.4) et (2.9) ? Dans le premier cas, on peut trouver des approximations suffisamment simples de l'équation pour lesquelles l'existence de solutions est claire, par exemple en discrétisant le temps à l'aide d'un schéma d'Euler (voir par exemple [9–11, 58, 65]). On déduira ensuite par un critère de tension la convergence de la suite de solutions approchées quand l'approximation tendra vers 0. Une identification des limites possibles montrera alors que le processus limite sera la solution recherchée.

Pour avoir des résultats d'existence ou d'unicité dans le cas non linéaire (2.9), une manière de procéder est de construire une approximation de l'équation non linéaire par une équation linéaire en dimension plus grande en utilisant un système de particules. Plus précisément, on va chercher à construire une estimation de la loi u_t de X_t, ce qui peut s'obtenir en considérant la loi empirique $\frac{1}{N}\sum_{n=1}^N \delta_{X^n}$ où les $(X_t^n)_{t\geq 0}$ sont des approximations de $(X_t)_{t\geq 0}$. Pour obtenir ces approximations du processus on considère la solution du système d'équations différentielles stochastiques obtenu en remplaçant chaque occurrence de u_t par son approximation $\frac{1}{N}\sum_{i=1}^n \delta_{X_t^i}$.

On obtient alors non plus une, mais un grand nombre N d'équations différentielles stochastiques qui cette fois-ci sont linéaires. Précisément, on obtient le système

$$\begin{cases} \mathrm{d}X_t^{n,N} = & \frac{1}{N}\sum_{m=1}^{N} b(X_t^{n,N}, X_t^{m,N})\mathrm{d}t + \frac{1}{N}\sum_{m=1}^{N} \sigma(X_t^{n,N}, X_t^{m,N})\mathrm{d}W_t^n, \quad \forall n \in \{1,\dots,N\} \\ (X_0^{n,N})_{n\in\{1,\dots,N\}} & \textit{i.i.d.} \text{ de loi } u_0 \end{cases},$$

où les $(W_t^n)_{t\geq 0}$ sont des mouvements Browniens indépendants issus de 0.

Le type de résultats que l'on souhaite ensuite obtenir sur ce genre de système sont des résultats de *propagation du chaos*, voir par exemple [64]. Précisément, il s'agit de montrer que pour un entier k fixé, la distribution jointe des k premières copies du système $(X^{1,N},\dots,X^{k,N})$ converge quand le nombre de particules N tend vers l'infini vers la distribution de k copies indépendantes du système.

2.2.2 Méthodes de réduction de variance pour les espérances de type Feynman-Kac

Si l'on veut utiliser l'écriture (2.8) pour simuler la solution de l'équation (2.7), une première idée est de simuler un grand nombre de trajectoires indépendantes $(\bar{X}_t^{x,n})_{t\geq 0, n=1,\dots,N}$ de même loi que $(X_t^x)_{t\geq 0}$ puis de calculer la moyenne empirique

$$\frac{1}{N}\sum_{n=1}^{N} f(\bar{X}_t^{x,n}) e^{-\int_0^t V(\bar{X}_s^{x,n})\mathrm{d}s}$$

pour avoir une estimation de (2.8), sur le principe de la méthode de Monte Carlo. Toutefois, l'efficacité de cette méthode dépend de la variance de la variable aléatoire à intégrer. Ici, le facteur exponentiel a en général une variance très grande, ce qui entraîne une mauvaise estimation de l'espérance.

Les systèmes de particules permettent de passer outre cette difficulté en évitant de calculer le poids exponentiel. Pour ce faire, on utilise une procédure de meurtre et duplication : les particules ayant une énergie importante vont être multipliées pendant la simulation, alors que celles ayant une énergie faible vont être tuées, de sorte que les zones de basse énergie, qui sont majoritaires dans le poids exponentiel, seront mieux explorées, voir par exemple [59].

Pour mettre en œuvre cette procédure, on écrit le potentiel V sous la forme de la différence de deux potentiels positifs V_m et V_d, les lettres d et m signifiant respectivement "duplication" et "meurtre" :

$$V = V_m - V_d.$$

Des choix naturels de telles décompositions sont

$$V_m = V - \sup(V), \quad V_d = 0$$

ou encore

$$V_m = V \vee 0, \quad V_d = (-V) \vee 0.$$

On simule ensuite des particules $(\tilde{X}_t^{x,n})_{t\geq 0, n=1,\dots,N}$ évoluant indépendamment selon des dynamiques Markoviennes de générateur $\mathcal{L}$. Chaque particule $\tilde{X}_t^{x,n}$, va, à taux $V_m(\tilde{X}_t^{x,n})$, être tuée et déplacée à l'emplacement d'une autre particule choisie uniformément parmi les autres particules. À l'opposé, chaque particule va, à taux $V_d(\tilde{X}_t^{x,n})$, être dupliquée en se voyant rejointe par une autre particule choisie indépendamment parmi les autres. Plus précisément, on définit τ comme le dernier instant de meurtre ou de duplication, ou comme $\tau = 0$, si ni l'un ni l'autre ne se sont encore produit. À chaque particule sont associées des variables indépendantes aléatoires E_n^d et E_n^m de loi exponentielle de paramètre 1. On définit les temps T_n^d et T_n^m par les formules

$$E_n^d = \int_\tau^{T_n^d} V_d(\tilde{X}_s^{x,n})\mathrm{d}s \quad \text{et} \quad E_n^m = \int_\tau^{T_n^m} V_m(\tilde{X}_s^{x,n})\mathrm{d}s.$$

Si $T_n^d \leq T_n^m$, alors la particule est tuée au temps T_n^m, sinon elle est dupliquée au temps T_n^d. Après chaque meurtre ou duplication, les variables aléatoires E_n^d et E_n^m sont réinitialisées pour chaque particule.

Un analogue en temps discret de la formule de Feynman-Kac est présenté dans [27], sous la forme $\mathbb{E}[F(X_K)\prod_{k=1}^K G(X_k)]$, où le processus à temps discret $(X_k)_{k\in\mathbb{N}}$ est une chaîne de Markov. Dans ce cadre, le terme produit joue le rôle du poids exponentiel, et va aussi causer une grande variance empêchant le calcul efficace de la moyenne par la méthode de Monte Carlo standard. En revanche, ce problème peut toujours être contourné par l'utilisation de systèmes de particules. On simulera ici des particules $(X_k^n)_{k\in\mathbb{N},n=1,\dots,N}$ évoluant selon la même dynamique que la chaîne $(X_k)_{k\in\mathbb{N}}$, mais qui seront rééchantillonnées à chaque pas de temps en donnant à chacune un poids proportionnel à $G(X_k^n)$. Plus précisément, pour chaque particule, on tire une variable aléatoire intermédiaire $\tilde{X}_{k+1}^n$ obtenue en faisant une transition Markovienne à partir de X_k^n. Ensuite, pour chaque $n = 1, \dots, N$, on tire X_k^n aléatoirement parmi les $(\tilde{X}_k^n)_{n=1,\dots,N}$, chacune ayant une probabilité $G(\tilde{X}_k^n)/\sum_{m=1}^N G(\tilde{X}_k^m)$ d'être tirée.

Dans chacun des deux cas ci-dessus, temps discret comme continu, on peut approcher l'espérance de Feynman-Kac en prenant la moyenne empirique du système de particules multipliée par un facteur de renormalisation. Il n'y a donc pas à calculer le terme exponentiel ou, en temps discret, le produit, réduisant de fait la variance de la variable aléatoire estimée.

3

Simulation moléculaire et applications

3.1 Quelques notions de physique statistique et de physique quantique

3.1.1 La physique statistique

La physique statistique a été développée dans le but d'étudier des systèmes physiques complexes, pour lesquelles une résolution par les méthodes exactes classiques s'avérait impossible. Typiquement, un volume macroscopique de gaz est composé d'un nombre de molécules de l'ordre du nombre d'Avogadro $\mathcal{N} = 6 \times 10^{23}$. Si l'on voulait décrire exactement le comportement du gaz à l'échelle microscopique, on obtiendrait un système d'équations différentielles ayant un nombre gigantesque de variables, qui serait impossible à résoudre directement, même en utilisant un ordinateur moderne.

La physique statistique, initiée par Ludwig Boltzmann à la fin du XIX$^{\text{e}}$ siècle, cherche à résoudre ce problème en adoptant une démarche probabiliste : au lieu de considérer indépendamment chaque particule du gaz et son équation d'évolution, on considère le gaz de manière globale, en faisant les calculs sur la distribution des positions et des vitesses des particules. L'équation d'évolution du système ne sera alors plus un système d'équations différentielles ordinaires décrivant le mouvement de chaque particule, mais une équation aux dérivées partielles décrivant l'évolution de la densité.

Formellement, la physique statistique va décrire un système physique par deux éléments : un espace d'état $\mathcal{E}$, qui désignera le plus souvent un ouvert de l'espace affine $\mathbb{R}^n$, et une fonction V dite d'*énergie*, définie sur $\mathcal{E}$ et à valeurs dans $\mathbb{R}$. L'espace $\mathcal{E}$ modélise l'ensemble de toutes les configurations possibles du système, alors que la fonction V associe à chaque configuration l'énergie correspondante.

Toutes les propriétés physiques macroscopiques du système sont alors représentées dans la mesure $\frac{1}{Z}e^{-\beta V(x)}\mathrm{d}x$, appelée *mesure de Gibbs*. On suppose que V et β sont tels que cette mesure soit finie, et on note $Z = \int_{\mathcal{E}} e^{-\beta V(x)}\mathrm{d}x$ la constante de normalisation, de sorte que la mesure de Gibbs soit une mesure de probabilité. Cette mesure minimise l'entropie par rapport à la mesure uniforme parmi les distributions de probabilité d'énergie fixée. Plus précisément, si on définit l'entropie relative de μ par rapport à $\mathrm{d}x$ par

$$\mathbf{Ent}(\mu) = \begin{cases} \int_{\mathcal{E}} f(x)\ln(f(x))\mathrm{d}x, & \text{avec } f = \frac{\mathrm{d}\mu}{\mathrm{d}x},\ \text{si } \mu \ll \mathrm{d}x, \\ 0 & \text{sinon} \end{cases},$$

alors $\frac{1}{Z}e^{-\beta V(x)}\mathrm{d}x$ vérifie

$$\mathbf{Ent}\left(\frac{1}{Z}e^{-\beta V(x)}\mathrm{d}x\right) = \inf\left\{\mathbf{Ent}(\nu), \nu \text{ probabilité}, \int_{\mathcal{E}} V(x)\mathrm{d}\nu(x) = \frac{1}{Z}\int_{\mathcal{E}} V(x)e^{-\beta V(x)}\mathrm{d}x\right\}.$$

L'entropie relative est une bonne façon de définir une distance entre deux mesures de probabilité, car on n'a $\mathbf{Ent}(\nu) = 0$ que dans le cas où $\nu(\mathrm{d}x) = \mathrm{d}x$, en raison de l'inégalité de Jensen appliquée

à la fonction convexe $x\ln(x)$. La mesure de Gibbs est donc au sens de l'entropie la mesure la plus proche de la mesure uniforme, à énergie fixée.

Le paramètre β apparaissant dans la définition de la mesure de Gibbs est homogène à l'inverse d'une température. Faire varier ce paramètre donne plusieurs mesures de Gibbs différentes, associées à des énergies totales différentes.

3.1.2 La physique quantique

La physique quantique décrit le comportement de la matière à l'échelle subatomique. L'idée directrice de la physique quantique est de ne plus voir les particules considérées comme des objets ponctuels ayant une masse et une position bien définies, mais en leur attribuant une probabilité de présence relativement étendue, aussi bien en position qu'en vitesse. Les physiciens désignent ce phénomène sous le nom d'indétermination.

On voit ici que l'aléa sur le système est d'une nature totalement différente en mécanique statistique et en mécanique quantique : alors qu'en mécanique statistique l'aléa n'est dû qu'à notre incapacité à connaître exactement l'état du système, en physique quantique, le système est *par nature* indéterminé.

L'état d'un système de particules physiques en mécanique quantique est représenté par une fonction φ, dite *fonction d'onde*, à valeur dans le corps $\mathbb{C}$ des nombres complexes. Pour un observateur, cette fonction d'onde a une interprétation probabiliste, puisque sur un grand nombre d'observations de particules dans un état décrit par une fonction d'onde φ, la répartition statistique des observations se fera selon le carré du module de la fonction d'onde $|\varphi|^2$.

Pour décrire l'évolution d'un système physique, on a donc besoin d'une équation décrivant l'évolution de la fonction d'onde. Cette équation fondamentale de la physique quantique est l'équation de Schrödinger, qui est le pendant quantique de la relation fondamentale de la dynamique de Newton en mécanique classique. Il s'agit de l'équation aux dérivées partielles suivante :

$$i\partial_t\varphi_t(x) = -\Delta\varphi_t(x) + V(x)\varphi_t(x), \quad (x,t) \in \mathbb{R}^d \times (0,\infty). \tag{3.1}$$

La fonction V désigne ici le champ de force auquel le système considéré est soumis.

Dans les cas physiques, le potentiel V sera tel que l'opérateur de Schrödinger, ou opérateur Hamiltonien $\Delta - V$ sera diagonalisable, avec un spectre discret et minoré. On peut par exemple penser, pour Ω un ouvert borné de $\mathbb{R}^n$, au cas $V(x) = 0$, si x est dans Ω et $V(x) = \infty$ sinon, correspondant à une particule confinée dans Ω. L'opérateur de Schrödinger est alors l'opérateur Laplacien sur un domaine borné, dont le caractère diagonalisable est un résultat classique d'analyse fonctionnelle.

Dans le cas diagonalisable, la résolution de (3.1) repose sur le calcul des différentes valeurs propres et fonctions propres associées de l'opérateur de Schrödinger, ce qui revient au calcul de la solution de l'*équation de Schrödinger stationnaire*

$$-\Delta\varphi(x) + V(x)\varphi(x) = E\varphi(x), \quad x \in \mathbb{R}^d \tag{3.2}$$

où les inconnues sont la fonction φ ainsi que la valeur E de l'énergie. En effet, si l'on décompose la fonction φ_0 sur la base de vecteurs propres $(\psi_k)_{k\geq 0}$ de $-\Delta + V$ sous la forme

$$\varphi_0 = \sum_{k=0}^{\infty} \lambda_k \psi_k,$$

alors la solution de l'équation de Schrödinger s'écrira

$$\varphi_t = \sum_{k=0}^{\infty} \lambda_k e^{itE_k} \psi_k,$$

où E_k est la valeur propre de $-\Delta + V$ associée à ψ_k.

C'est pour cela qu'une des questions fondamentales en physique quantique est de calculer non pas la solution de l'équation de Schrödinger de manière générale, ce qui serait trop complexe, mais de calculer les énergies propres du système, à savoir les valeurs propres de l'opérateur de Schrödinger, et les états propres associés. Notamment, l'énergie fondamentale du système, c'est-à-dire la plus petite valeur propre de l'opérateur de Schrödinger, et son vecteur propre associé jouent un rôle particulièrement important, car ils correspondent à l'état non excité du système.

3.2 Calculs d'énergies libres par la méthode de la force biaisante adaptative

Un problème ayant de nombreuses applications en physique statistique est le problème du calcul d'énergies libres. Étant donné un système, ce qui en physique statistique est la donnée d'un espace d'état $\mathcal{E}$ et d'une fonction d'énergie V, on peut chercher à étudier non pas le comportement global du système, mais seulement le comportement de certaines quantités d'intérêt. Plus précisément, on considère une fonction ξ, appelée *coordonnée de réaction* définie sur $\mathcal{E}$ à valeurs dans un espace de petite dimension, et on ne s'intéresse pas à l'évolution du système dans $\mathcal{E}$, mais plutôt à l'évolution de l'image par ξ du système. La fonction ξ peut donc être comprise comme une observable macroscopique.

On peut penser comme exemple à l'évolution d'une réaction chimique. Le système considéré est un ensemble de molécules qui vont réagir les unes avec les autres. L'information à retenir n'est pas alors la position microscopique de chaque molécule, mais plutôt l'avancement global de la réaction, quantifié par exemple par un réel de $[0, 1]$ correspondant à la proportion de molécules ayant réagi. Un autre exemple est la conformation des protéines. Une protéine est une longue chaîne d'acides aminés reliés les uns avec les autres d'une certaine manière, de sorte à donner une forme particulière à la protéine. Étudier la forme des protéines est utile pour pouvoir comprendre les mécanismes d'action de celles-ci. Ce que l'on veut alors retenir dans une configuration particulière de la protéine n'est pas la position exacte de chaque acide aminé, mais plutôt une forme générale, par exemple l'écartement entre les deux extrémités, ou l'angle formé par deux sous-ensembles particuliers d'acides aminés.

Une information importante pour étudier cette quantité d'intérêt, est la "trace" de l'énergie globale du système correspondant à cette coordonnée, appelée énergie libre. Comme en physique statistique l'énergie n'intervient qu'à travers la mesure de Gibbs, il est naturel de définir l'énergie libre comme la fonction A telle que la mesure $e^{-\beta A(z)}\mathrm{d}z$ soit la mesure image de la mesure $\frac{1}{Z}e^{-\beta V(x)}\mathrm{d}x$ par l'application ξ. On peut alors donner une expression de l'énergie libre, grâce à la *formule de la co-aire* qui est une généralisation du théorème de Fubini à des coordonnées qui ne sont pas orthogonales. On obtient

$$A'(z) = \mathbb{E}[F(X)|\xi(X) = z], \tag{3.3}$$

où X est tiré selon la mesure de Gibbs, et F est la fonction définie par

$$F = \frac{\nabla\xi \cdot \nabla V}{|\nabla\xi|^2} - \frac{1}{\beta}\mathrm{div}\left(\frac{\nabla\xi}{|\nabla\xi|^2}\right).$$

La quantité $-F(x)$ est à comprendre comme la composante selon ξ de la force s'exerçant au point x, et $-A'(z)$ est donc la force moyenne à l'équilibre, sachant que le système s'envoie sur z par la fonction ξ. La fonction A peut donc bien être assimilée à une énergie potentielle sur la composante macroscopique du système.

Pour calculer l'énergie libre, une méthode naïve serait de simuler une trajectoire ergodique par rapport à la mesure de Gibbs $\frac{1}{Z}e^{-\beta V(x)\mathrm{d}x}$ jusqu'à avoir un échantillon important de variables

aléatoires de loi $\frac{1}{Z}e^{-\beta V(x)}\mathrm{d}x$, puis de prendre l'image par ξ de cet échantillon. Une dynamique typique permettant ce genre d'échantillonnage est la dynamique de Langevin sur-amortie,

$$\mathrm{d}X_t = -\nabla V(X_t)\mathrm{d}t + \sqrt{\frac{2}{\beta}}\mathrm{d}W_t. \tag{3.4}$$

Bien entendu, cette méthode serait trop lente à mettre en place, car les problèmes physiques ou biologiques typiques sont par nature *métastables*, c'est-à-dire que bien que la trajectoire soit ergodique, la convergence est très lente car le système reste bloqué pendant de longue périodes dans des minima locaux de la fonction d'énergie V.

La méthode de la force biaisante adaptative, introduite dans [25, 32], permet de calculer cette énergie libre de manière bien plus efficace, en se débarrassant de certains états métastables. L'idée est de rajouter un terme à l'équation de Langevin (3.4) qui va éloigner le système des états métastables déjà visités. Remarquons que l'expression (3.3) est toujours valable si la variable X est distribuée non pas selon la mesure de Gibbs, mais selon la mesure modifiée

$$\frac{1}{Z}e^{-\beta(V(x)-A(\xi(x)))}\mathrm{d}x, \tag{3.5}$$

puisque l'expression de A' est prise conditionnellement à $\xi(X)$ et que la loi de X n'est modifiée que par un facteur ne dépendant que de $\xi(x)$.

On peut donc obtenir la même expression de A' en échantillonnant X grâce à la dynamique

$$\mathrm{d}X_t = -\nabla(V(X_t) - A(\xi(X_t)))\mathrm{d}t + \sqrt{2/\beta}\mathrm{d}W_t, \tag{3.6}$$

dont la mesure invariante est donnée par (3.5). Bien évidemment cette méthode ne peut pas être mise en place directement, puisque la dynamique (3.6) dépend de la fonction A qui est précisément ce que l'on cherche à calculer. Cependant, on peut calculer une approximation de la fonction A en utilisant l'expression (3.3). Une dynamique possible est alors

$$\begin{cases} \mathrm{d}X_t &= -\nabla(V(X_t) - A_t(\xi(X_t)))\mathrm{d}t + \sqrt{2/\beta}\mathrm{d}W_t \\ A'_t(z) &= \mathbb{E}[F(X_t)|\xi(X_t) = z] \end{cases}. \tag{3.7}$$

Dans [48] il est montré, sous certaines hypothèses, que toute solution de cette équation permet bien, sous réserve d'existence, d'obtenir A' comme limite en temps long de A'_t. La preuve utilise un découpage de l'espace dans deux directions, une selon ξ, correspondant à une évolution macroscopique, et une orthogonale à ξ, correspondant à une évolution microscopique. Dans [48], on fait l'hypothèse que les mesures de Gibbs conditionnées à une valeur de ξ constante vérifient une inégalité de Sobolev logarithmique (voir par exemple [6,61]), ce qui signifie une convergence rapide au niveau microscopique. En revanche, au niveau macroscopique, la vitesse de convergence rapide est assurée, puisque la densité de $\xi(X_t)$ satisfait l'équation de la chaleur, et ce même si $\xi(X_t)$ ne suit pas une dynamique Brownienne. D'une certaine manière, l'ajout d'un terme de biais à la dynamique de Langevin permet de tuer les métastabilités au niveau macroscopique.

La simulation du système (3.7) est une question difficile, car elle nécessite d'estimer une espérance conditionnelle. Pour ce faire, les praticiens utilisent généralement des moyennes ergodiques en temps long. Toutefois, il parait naturel de simuler un tel système par une méthode de Monte Carlo, en utilisant un système de particules. En effet on peut construire une estimation de l'espérance conditionnelle grâce à un estimateur de Nadaraya-Watson, de la forme

$$\mathbb{E}[F(X)|\xi(X) = z] \simeq \frac{\frac{1}{N}\sum_{i=1}^{N}\varphi_{\alpha,\varepsilon}(\xi(X_i) - z)F(X_i)}{\frac{1}{N}\sum_{i=1}^{N}\varphi_{\alpha,\varepsilon}(\xi(X_i) - z)}.$$

Dans cette expression, $\varphi_{\alpha,\varepsilon}$ est une approximation de la masse de Dirac, qui est choisie de la forme $\varphi_{\alpha,\varepsilon} = \alpha + \psi_\varepsilon$, où ψ_ε est une fonction positive d'intégrale unité à support dans $[-\varepsilon, \varepsilon]$. On peut donc définir le système de particules suivant pour approcher la solution de (3.7) :

$$\begin{cases} \mathrm{d}X^{\alpha,\varepsilon}_{t,n,N} = & \left(-\nabla V(X^{\alpha,\varepsilon}_{t,n,N}) + \dfrac{\sum_{m=1}^N \varphi_\eta(X^{\alpha,\varepsilon,1}_{t,n,N} - X^{\alpha,\varepsilon,1}_{t,m,N}) F(X^{\alpha,\varepsilon}_{t,m,N})}{\sum_{m=1}^N \varphi_\eta(X^{\alpha,\varepsilon,1}_{t,n,N} - X^{\alpha,\varepsilon,1}_{t,m,N})} \nabla\xi(X^{\alpha,\varepsilon}_{t,n,N})\right)\mathrm{d}t + \sqrt{2}\mathrm{d}W^n_t, \\ X^{\alpha,\varepsilon}_{0,n,N} \qquad \text{de loi donnée} \end{cases},$$

pour $n = 1, \ldots, N$. Sous certaines hypothèse sur l'énergie potentielle V, nous avons montré que l'on a existence et unicité d'une solution faible de l'équation (3.7), et que l'approximation particulaire de A' converge vers la vraie énergie libre A'_t au temps t quand le nombre de particules tend vers l'infini et que le noyau régularisant pour le calcul de l'espérance conditionnelle tend vers une masse de Dirac. Plus précisément :

$$\mathbb{E}\left[\int_0^T \left\| \frac{\sum_{n=1}^N F(X^{\alpha,\varepsilon}_{t,n,N})\varphi_\eta(. - X^{\alpha,\varepsilon,1}_{t,n,N})}{\sum_{n=1}^N \varphi_\eta(. - X^{\alpha,\varepsilon,1}_{t,n,N})} - A'_t \right\|_{\mathbb{L}^\infty(\mathbb{T})} \mathrm{d}t\right] = \mathcal{O}\left(\alpha + \sqrt{\varepsilon} + \frac{1}{\sqrt{N}} e^{\frac{K}{\alpha\varepsilon^2}}\right).$$

3.3 Calculs de sensibilité en chimie quantique

Le but de la chimie quantique est de calculer les structures électroniques de molécules. Dans le cadre de la physique quantique, cela signifie calculer la fonction d'onde du système composé des noyaux et des électrons de la molécule.

Ce genre de problème amène à la résolution d'équations aux dérivées partielles en très grande dimension, ce qui fait qu'une résolution par des méthodes numériques classiques est très ardue. Par exemple, pour la molécule Li_8, la dimension du problème est $8 \times (3+1) \times 3 = 96$, puisque chacun des huit atomes de lithium est composé d'un noyau et de trois électrons évoluant en trois dimensions.

Une première simplification du problème est l'approximation de Born-Oppenheimer. Cette approximation consiste à considérer que les deux types de particules, électrons et noyaux, ont des vitesses d'évolution caractéristiques très différentes ; pour être précis, les électrons se déplacent beaucoup plus rapidement que les noyaux. On va donc séparer le problème en deux parties, et calculer différemment la position des électrons et celles des noyaux. Pour calculer l'état fondamental des électrons, on suppose alors que les noyaux sont fixés dans l'espace, ce qui induit un champ d'énergie potentielle V dans lequel évoluent les électrons. Typiquement, la fonction V est de la forme

$$V(x) = \sum_{i=1}^N V_1(x_i) + \sum_{1 \le i < j \le N} V_2(x_i - x_j),$$

où V_1 désigne le champ de force que l'ensemble des noyaux applique sur chaque électron, et V_2 désigne la force d'interaction entre deux électrons. V_1 sera généralement de la forme

$$V_1(x) = -\sum_{k=1}^K z_k \rho_k * \frac{1}{|x|},$$

où z_k est la charge du $k^{\text{ième}}$ noyau, et ρ_k représente l'étendue spatiale du noyau. Par exemple, on pourra avoir $\rho_k = \delta_{y_k}$ pour un noyau ponctuel en position y_k, ou bien ρ_k pourra être une fonction régulière à support dans un petit voisinage de y_k, pour un noyau diffus.

Dans le cadre de l'approximation de Born-Oppenheiemer, le système étudié du point de vue quantique est composé uniquement d'électrons, qui sont des particules fermioniques, cette pro-

priété se traduisant par le fait que la fonction d'onde sera à chercher dans l'espace des fonctions antisymétriques.

Pour calculer la bonne position des noyaux, on suppose qu'ils suivent une dynamique classique. L'énergie globale du système est donnée par

$$\mathcal{V}(x_1,\dots,x_K) = E_0(x_1,\dots,x_K) + \sum_{1\le k<q\le K} \frac{z_k z_q}{|y_k - y_q|}, \tag{3.8}$$

où E_0 désigne l'énergie fondamentale des électrons. L'évolution des positions des noyaux est donc régie par l'équation de Newton

$$\partial_t^2(m_1 x_1,\dots,m_K x_K) = -\nabla\mathcal{V}(x_1,\dots,x_K),$$

où m_k est la masse du $k^{\text{ième}}$ noyau. La position d'équilibre des noyaux peut donc être calculée comme la position minimisant la fonction $\mathcal{V}$. Un problème important est donc de savoir calculer la force $\nabla\mathcal{V}(x_1,\dots,x_K)$ s'exerçant sur les noyaux.

3.3.1 Les méthodes de Monte Carlo en chimie quantique

En chimie quantique, les calculs de valeurs propres d'opérateurs de Schrödinger sont généralement difficiles à traiter directement, à cause de leur grande dimension. C'est pourquoi un certain nombre de méthodes probabilistes ont été développées pour résoudre ces problèmes. Ces méthodes sont désignées collectivement sous le nom générique de méthodes de Monte Carlo quantiques.

L'énergie fondamentale d'un opérateur de Schrödinger H est définie précisément par

$$E = \inf\{\langle\varphi, H\varphi\rangle\,,\ \varphi\in\mathcal{D}(\mathcal{H}),\ \langle\varphi,\varphi\rangle = 1\}, \tag{3.9}$$

où $\mathcal{H}$ est un espace de Hilbert contenu dans $\mathbb{L}^2(\Omega)$ pour un ouvert Ω de $\mathbb{R}^d$, et $\mathcal{D}(q_H)$ est le domaine de la forme quadratique $\langle\varphi, H\varphi\rangle$. La borne inférieure dans l'expression (3.9) est atteinte pour une fonction normalisée ψ.

Nous allons nous intéresser à deux méthodes probabilistes différentes pour traiter ces problèmes : la méthode de Monte Carlo variationnelle et la méthode de Monte Carlo diffusive. Ces deux méthodes se servent d'une interprétation de E comme un objet probabiliste. Supposons que l'on connaisse, à une constante multiplicative près, une bonne approximation φ, dite *fonction d'essai*, de l'état fondamental ψ. On peut écrire,

$$E = \langle\psi, H\psi\rangle \simeq \frac{\langle\varphi, H\varphi\rangle}{\langle\varphi,\varphi\rangle} = \mathbb{E}\left[\frac{H\varphi}{\varphi}(X)\right], \tag{3.10}$$

où X est une variable aléatoire dont la loi est donnée par la fonction $\frac{\varphi^2}{\langle\varphi,\varphi\rangle}$. La variable aléatoire $\frac{H\varphi}{\varphi}(X)$ peut donc servir d'estimateur pour calculer l'énergie fondamentale. De plus cette quantité est une surestimation de l'énergie fondamentale, puisque E est la plus petite des valeurs propres. Cette propriété est importante, car elle permet d'estimer l'erreur commise sur le calcul de E.

Pour pouvoir utiliser ce résultat, il faut être capable de simuler des variables aléatoires de loi $\varphi^2/\langle\varphi,\varphi\rangle$.Cela est possible en utilisant une dynamique de Langevin sur-amortie

$$\mathrm{d}X_t = -\frac{1}{2}\frac{\nabla\varphi}{\varphi}(X_t)\mathrm{d}t + \mathrm{d}W_t.$$

En effet, sous de bonnes hypothèses sur φ, cette dynamique est ergodique de mesure invariante $\varphi^2/\langle\varphi,\varphi\rangle$ On peut donc estimer l'espérance (3.10), par exemple en calculant des moyennes ergodiques sur les trajectoires de X_t.

Dans la méthode de Monte Carlo diffusive, on va utiliser la formule

$$E = \frac{\langle \varphi, H\psi \rangle}{\langle \varphi, \psi \rangle} = \mathbb{E}\left[\frac{H\varphi}{\varphi}(X)\right],$$

où X est une variable aléatoire dont la loi est donnée par la fonction $\varphi\psi/\langle \varphi, \psi \rangle$. Cet estimateur de l'énergie fondamentale est également d'ordre élevé par rapport à l'erreur commise entre l'état fondamental ψ et son approximation $\varphi/\langle \varphi, \psi \rangle$.

Il est également possible de simuler des variables aléatoires de loi $\varphi\psi/\langle \varphi, \psi \rangle$ grâce à la méthode de Monte Carlo diffusive, qui est une méthode d'échantillonage préférentiel. Cette méthode est basée sur le fait que la solution de l'équation

$$\begin{cases} \partial_t \Phi_t(x) = -H\Phi_t(x), & (t,x) \in [0,\infty) \times \Omega, \\ \Phi_0 = \varphi, \end{cases}$$

est équivalente en temps long à la fonction $e^{-Et}\psi \langle \psi, \varphi \rangle$ pourvu que $\langle \psi, \varphi \rangle$ soit non nul. L'état fondamental de H peut donc s'exprimer sous la forme

$$E = \frac{\langle \psi, H\varphi \rangle}{\langle \psi, \varphi \rangle} = \lim_{t\to\infty} \frac{\langle \Phi_t, H\varphi \rangle}{\langle \Phi_t, \varphi \rangle}. \tag{3.11}$$

La fonction Φ_t est calculable par des méthodes probabilistes en utilisant la formule de Feynman-Kac suivante

$$\Phi_t(x) = \mathbb{E}\left[\varphi(x + W_t) e^{-\int_0^t V(x+W_s)\mathrm{d}s}\right],$$

où $(W_t)_{t\geq 0}$ est un mouvement Brownien standard. Il est toutefois préferable de ne pas utiliser directement cette formulation et de passer par une méthode d'échantillonnage préférentiel, le facteur exponentiel pouvant être la cause d'une grande variance. On utilise donc plutôt en pratique la fonction $\tilde{f}_t(x) = \varphi(x)\Phi_t(x)/\langle \varphi, \varphi \rangle$ qui nous permet de donner une autre expression du membre de droite de (3.11) :

$$\frac{\langle \Phi_t, H\varphi \rangle}{\langle \Phi_t, \varphi \rangle} = \frac{\int_\Omega \frac{H\varphi}{\varphi}(x)\tilde{f}_t(x)\mathrm{d}x}{\int_\Omega \tilde{f}_t(x)\mathrm{d}x}.$$

La fonction $\tilde{f}$ est solution de

$$\begin{cases} \partial_t \tilde{f}(x) & = \frac{1}{2}\Delta\tilde{f}(x) + \nabla \cdot \left(\frac{\nabla\varphi}{\varphi}(x)\tilde{f}(x)\right) - \frac{H\varphi(x)}{\varphi(x)}\tilde{f}(x), \quad (t,x) \in [0,\infty) \times \Omega \\ f_0(x) & = \varphi^2(x)/\langle \varphi, \varphi \rangle, \quad x \in \Omega \end{cases}.$$

Cette équation aux dérivées partielles s'interprète comme l'équation vérifiée par la fonction h définie par

$$\int_\Omega g(x)h_t(x)\mathrm{d}x = \mathbb{E}\left[g(X_t)e^{-\int_0^t \frac{H\varphi}{\varphi}(X_s)\mathrm{d}s}\right],$$

où $(X_t)_{t\geq 0}$ suit la dynamique

$$\begin{cases} \mathrm{d}X_t & = -\frac{\nabla\varphi}{\varphi}(X_t)\mathrm{d}t + \mathrm{d}W_t \\ X_0 & \text{de loi } \varphi^2/\langle \varphi, \varphi \rangle \end{cases}.$$

Toutefois, il se trouve que les fonctions $\tilde{f}$ et h sont en général distinctes, sauf si la fonction d'essai φ a le même ensemble de zéros, appelé *ensemble des nœuds* en physique, que le véritable état fondamental ψ. Il est quand même possible de calculer la quantité

$$E_{DMC}(t) = \frac{\int_\Omega \frac{H\varphi}{\varphi}(x)h_t(x)\mathrm{d}x}{\int_\Omega h_t(x)\mathrm{d}x} = \frac{\mathbb{E}\left[\frac{H\varphi}{\varphi}(X_t)e^{-\int_0^t \frac{H\varphi}{\varphi}(X_s)\mathrm{d}s}\right]}{\mathbb{E}\left[e^{-\int_0^t \frac{H\varphi}{\varphi}(X_s)\mathrm{d}s}\right]}$$

qui est une approximation de $E(t)$. En fait, il est montré dans [21] que, sous de bonnes hypothèses, la quantité $E_{DMC}(t)$ converge en temps long vers une constante E_{DMC}, qui est une surestimation de E pouvant aussi s'exprimer comme la solution d'un problème variationnel. Plus précisément, E_{DMC} est l'énergie fondamentale de l'opérateur H restreint à une composante connexe de $\Omega \setminus \varphi^{-1}(0)$. Cette erreur systématique dûe au choix de φ est connue sous le nom d'*approximation de nœuds fixes.*

3.3.2 Estimateurs zéro biais/zéro variance

Il peut également être intéressant, par exemple dans le cas du calcul des forces dans l'approximation de Born-Oppenheimer, d'estimer la dérivée $\partial_\lambda^0 E_\lambda$ de l'énergie fondamentale d'un opérateur de Schrödinger H_λ dépendant d'un paramètre λ (pour une fonction f_λ dépendant d'un paramètre, on notera par $\partial_\lambda^0 f_\lambda$ la dérivée en 0 de f par rapport à λ). En effet, si l'on pose $H_\lambda = -\Delta + V_\lambda$, où V_λ est l'énergie potentielle créée par une certaine position des noyaux indexée par un paramètre λ, le calcul de $\partial_\lambda^0 E_\lambda$ intervient dans le calcul des forces (3.8).

Cela peut également permettre de calculer la moyenne d'une observable, puisqu'en posant $H_\lambda = H + \lambda \mathbf{O}$ pour un opérateur $\mathbf{O}$, la dérivée de l'énergie fondamentale par rapport à λ est égale à la moyenne de l'observable $\mathbf{O}$. En effet, si ψ_λ désigne le vecteur propre normalisé associé à E_λ, on a

$$\begin{aligned}\partial_\lambda^0 E_\lambda = \partial_\lambda^0 \langle \psi_\lambda, H_\lambda \psi_\lambda \rangle &= 2 \langle \partial_\lambda^0 \psi_\lambda, H\psi \rangle + \langle \psi, \partial_\lambda^0 H_\lambda \psi \rangle \\ &= 2E \langle \partial_\lambda^0 \psi_\lambda, \psi \rangle + \langle \psi, \mathbf{O}\psi \rangle \\ &= \langle \psi, \mathbf{O}\psi \rangle .\end{aligned}$$

Les estimateurs présentés dans la partie 3.3.1 sont des estimateurs dont le biais et la variance sont très petits si φ est choisie proche de ψ, puisque la variable aléatoire $\frac{H\varphi}{\varphi}(X)$ devient une constante égale à la valeur E à estimer quand $\varphi = \psi$. En fait on peut montrer que le biais et la variance sont de l'ordre de $\delta\psi^2$ dans une certaine métrique, où

$$\delta\psi = \varphi / \sqrt{\langle \varphi, \varphi \rangle} - \psi$$

dans le cas de la méthode de Monte Carlo variationnelle, et

$$\delta\psi = \varphi / \langle \varphi, \psi \rangle - \psi$$

dans le cas de la méthode de Monte Carlo diffusive.

Si l'on veut généraliser ces estimateurs à petit biais et petite variance au cas de la dérivée en λ de l'énergie il semble naturel d'utiliser, par exemple pour la méthode Monte Carlo variationnelle, l'égalité

$$\partial_\lambda^0 E_\lambda = \partial_\lambda^0 \left(\frac{\langle \varphi, H_\lambda \varphi \rangle}{\langle \varphi, \varphi \rangle} \right) = \mathbb{E} \left[\frac{\partial_\lambda^0 H_\lambda \varphi}{\varphi}(X) \right],$$

où X a pour loi $\varphi^2 / \langle \varphi, \varphi \rangle$. Cependant l'estimateur obtenu ne vérifie pas la propriété d'ordre élevé pour le biais et la variance, notamment car il ne devient en général pas constant si $\varphi = \psi$.

Une idée pour trouver des estimateurs à petit biais et petite variance, introduite dans [7, 68] est de choisir une fonction d'essai dépendante de λ. L'erreur s'exprimera alors à la fois en fonction de l'erreur entre φ et ψ, mais aussi de l'erreur entre $\partial_\lambda^0 \varphi_\lambda$ et $\partial_\lambda^0 \psi_\lambda$. On trouve alors les bons estimateurs en exprimant les dérivées

$$\partial_\lambda^0 E_\lambda = \partial_\lambda^0 \left(\frac{\langle \varphi_\lambda, H_\lambda \varphi_\lambda \rangle}{\langle \varphi_\lambda, \varphi_\lambda \rangle} \right) \quad \text{et} \quad \partial_\lambda^0 E_\lambda = \partial_\lambda^0 \left(\frac{\langle \varphi_\lambda, H_\lambda \psi_\lambda \rangle}{\langle \varphi_\lambda, \psi_\lambda \rangle} \right)$$

comme des espérances relativement à $\varphi^2/\langle\varphi,\varphi,\rangle$ et $\varphi\psi/\langle\varphi,\psi\rangle$. L'estimateur intéressant ici est l'estimateur obtenu par la méthode de Monte Carlo diffusive qui est plus précis, car non biaisé. Le problème est que l'estimateur obtenu n'est en fait pas calculable, puisqu'il dépend de la fonction ψ_λ qui est inconnue. Une manière d'échapper à ce problème est d'utiliser une diffusion dépendant de λ. Plus précisément, si l'opérateur de Schrödinger H_λ est de la forme $-\Delta + V_\lambda$, au lieu d'appliquer un estimateur précis mais non calculable à une trajectoire de dynamique de Monte Carlo diffusive, on dérive l'égalité

$$E_\lambda = \frac{\langle\psi_\lambda, H_\lambda\varphi_\lambda\rangle}{\langle\psi_\lambda,\varphi_\lambda\rangle} = \lim_{t\to\infty} \frac{\mathbb{E}\left[\frac{H_\lambda\varphi_\lambda}{\varphi_\lambda}(X_t^\lambda)e^{-\int_0^t \frac{H_\lambda\varphi_\lambda}{\varphi_\lambda}(X_s^\lambda)\mathrm{d}s}\right]}{\mathbb{E}\left[e^{-\int_0^t \frac{H_\lambda\varphi_\lambda}{\varphi_\lambda}(X_s^\lambda)\mathrm{d}s}\right]}, \tag{3.12}$$

pour obtenir une approximation, en temps long de $\partial_\lambda^0 E_\lambda$.

Dans (3.12), X_t^λ suit une dynamique de la forme

$$\mathrm{d}X_t^\lambda = -\nabla V_\lambda(X_t^\lambda)\mathrm{d}t + \mathrm{d}W_t.$$

La dérivation de (3.12) par rapport à λ fait donc apparaître des termes dépendants de la dérivée en λ de T_t de X_t^λ, appelée *vecteur tangent* qui vérifie l'équation différentielle ordinaire à coefficients aléatoires

$$\partial_t T_t = -\nabla^2 V_0(X_t^0)\cdot T_t - \partial_\lambda^0 \nabla V_\lambda(X_t^0).$$

Le premier terme du membre de droite va créer des accroissements exponentiels de la norme de T_t au voisinage des maxima et des points selles du potentiel V_0. Le vecteur T_t est donc un facteur multiplicatif pouvant prendre de grandes valeurs, faisant augmenter la variance.

Il est possible de réduire la variance de T_t en utilisant le formalisme des formules de Feynman-Kac : on va simuler des copies du processus (X_t, T_t) que l'on rééchantillonnera à intervalles réguliers de sorte à multiplier les copies dont le vecteur tangent prend de grandes valeurs.

4

Schéma numérique pour une loi de conservation scalaire fractionnaire

4.1 Lois de conservation

On appelle *loi de conservation* un système d'équations aux dérivées partielles de la forme

$$\partial_t u_t^i(x) + \sum_{j=1}^{d} \partial_{x_j} \left(f_t^{ij} \left(u_t^1(x), \dots, u^I(x) \right) \right) = g_t^i(x), \ i = 1, \dots, I, \quad (x,t) \in \mathbb{R}^d \times (0,\infty). \tag{4.1}$$

Ce type d'équation traduit, sous forme différentielle, le bilan d'une certaine grandeur extensive physique, comme par exemple, la masse, l'énergie, ou encore l'entropie. La quantité $u_t^i(x)$ correspond à la quantité de grandeur extensive de type i au point x à l'instant t.

Le deuxième terme du membre de gauche de (4.1) traduit le flux de la grandeur i, alors que le membre de droite est un terme source, correspondant à la création (ou à la destruction) de cette grandeur extensive, au taux $g_t^i(x)$ au point x. Pour un panorama sur les lois de conservation, voir par exemple [62].

4.1.1 Notions de solutions pour les lois de conservation scalaires

On considère dans cette partie des lois de conservation scalaire, c'est-à-dire pour lesquelles l'espace de départ est de dimension 1. Pour simplifier l'analyse, on considère une équation sans terme source. On obtient donc l'équation

$$\begin{cases} \partial_t v_t(x) + \partial_x (A(v_t(x))) = 0, \ (x,t) \in \mathbb{R} \times (0,\infty) \\ v_0 \text{ donné} \end{cases}. \tag{4.2}$$

On suppose par la suite que la condition initiale v_0 est bornée. Cette équation aux dérivées partielles est remarquable par le fait qu'elle peut admettre des solutions irrégulières, même si sa condition initiale est très lisse.

La méthode naïve pour résoudre l'équation (4.2) est de chercher des solutions régulières. Pour ce faire, on cherche, dans l'espace-temps $\mathbb{R} \times (0,\infty)$, des chemins, appelés caractéristiques, le long desquels la solution est constante, correspondant à des lignes de propagation de l'information contenue dans la solution. Concrètement, étant donnée une solution régulière v de la loi de conservation (4.2), on cherche une fonction γ_t satisfaisant :

$$v_t(\gamma_t) = v_0(\gamma_0),$$

pour tout temps t. En dérivant cette équation par rapport au temps, on obtient l'équation suivante,

$$\partial_t \gamma_t = A'(v_t(\gamma_t)). \tag{4.3}$$

Comme v est constante le long des trajectoires de γ, cette équation peut se récrire

$$\partial_t \gamma_t = A'(v_0(\gamma_0)).$$

Ainsi, la caractéristique issue de x à l'instant 0 a pour équation $\gamma_t = x + tA'(v_0(x))$.

L'équation (4.2) est alors entièrement résolue si les différentes caractéristiques constituent une partition de l'espace temps $\mathbb{R} \times (0, \infty)$: il suffit d'attribuer à $v_t(x)$ la valeur $v_0(x_0)$, où x_0 est le point d'origine de la caractéristique passant par le point (x, t).

Cette approche ne donne plus de résultat lorsque les caractéristiques peuvent se couper, où lorsqu'elles ne remplissent pas l'espace. Cette situation est illustrée par les deux exemples suivants : pour $A(x) = x^2/2$, considérons les conditions initiales $v_0(x) = \mathbf{1}_{(0,\infty)}(x)$ et $v_0(x) = \mathbf{1}_{(-\infty,0)}$. Dans le premier cas, les caractéristiques se répartissent en deux catégories : pour $x_0 < 0$ elles, ont pour équation $x = x_0$, et pour $x_0 > 0$, l'équation est $x = x_0 + t$. On voit alors qu'aucune caractéristique ne vient occuper le domaine $\{0 < x < t\}$. Ce phénomène est l'expression d'une onde de détente régularisante. Dans le deuxième cas, en revanche, les caractéristiques ont pour équation $x = x_0$ si $x_0 > 0$ ou $x = x_0 + t$ si $x < 0$, et par chaque point du même domaine $\{0 < x < t\}$ passent alors deux caractéristiques. Ici, on observe la propagation d'une onde de choc.

On doit donc chercher une autre méthode pour résoudre cette équation. On voit que ce phénomène de croisement des caractéristiques entraîne la non existence de solutions régulières. Il faut donc affaiblir la notion de solution pour pouvoir chercher des solutions à l'équation dans un espace plus étendu.

Une manière d'affaiblir la notion de solution est d'utiliser la théorie des distributions, c'est-à-dire que l'on teste la validité de la solution en intégrant par parties contre une fonction test régulière : une *solution faible* sera alors une fonction v de $\mathbb{L}^\infty(\mathbb{R} \times (0, \infty))$ (par exemple) telle que pour toute fonction φ régulière à support compact, on a

$$\int_{\mathbb{R}} v_0 \varphi_0 + \int_0^\infty \int_{\mathbb{R}} A(v_t) \partial_x \varphi_t \mathrm{d}t = 0.$$

Toutefois, il se trouve que la notion de solution a été ici trop affaiblie, et que l'on perd l'unicité de la solution. Il faut donc utiliser une notion de solution intermédiaire entre solutions faibles et fortes.

La "bonne" notion de solution pour la loi de conservation est la notion de *solution entropique* introduite en 1970 par Kruzhkov dans [44]. L'idée ici est de rajouter un petit terme de diffusion à l'équation, rendant ainsi la dynamique sous-jacente irréversible. Le terme de diffusion ajouté ayant un effet régularisant, l'équation obtenue aura de bonnes propriétés mathématiques. On fera alors tendre le terme de diffusion vers 0, et la suite de solutions de l'équation régularisée convergera vers une solution de l'équation (4.2), qui sera la "bonne", au sens où elle retranscrira bien les propriétés physiques de l'équation.

Plus précisément, l'équation

$$\partial_t v^\sigma(x) = \sigma \Delta v_t^\sigma(x) - \partial_x(A(v_t^\sigma(x))), \quad (x, t) \in \mathbb{R} \times (0, \infty) \tag{4.4}$$

a une unique solution v^σ, qui appartient à l'espace de Lebesgue à poids $\mathbb{L}^\infty((0, \infty), \mathbb{L}^1((1+x^2)^{-1}))$ et qui converge dans ce même espace, quand σ tends vers 0, vers une limite v qui se trouve être une solution faible de (4.2). Il faut donc trouver un critère permettant de caractériser cette solution parmi toutes les autres solutions faibles de (4.2). Si l'on multiplie l'équation (4.4) par $\eta'(v_t^\sigma)$, où η est une fonction convexe, on obtient formellement

$$\partial_t(\eta(v_t^\sigma(x))) + \partial_x(\psi(v_t^\sigma(x))) = \sigma \eta'(v_t^\sigma(x)) \Delta(v_t^\sigma(x)),$$

où la fonction ψ est une primitive de $A'\eta'$. En intégrant en espace et en temps contre une fonction test régulière à support compacte, et en intégrant par parties de sorte à faire agir les dérivations

sur la fonction test, on obtient la formulation intégrale :

$$\begin{aligned}&\int_{\mathbb{R}} \eta(v_0^\sigma)v^\sigma\varphi_0 + \int_0^\infty \int_{\mathbb{R}} \eta(v_t^\sigma)\partial_t v^\sigma\varphi_t + \psi(v_t^\sigma)\partial_x v^\sigma\varphi_t \mathrm{d}t + \sigma\int_0^\infty \int_{\mathbb{R}} \eta(v_t^\sigma)\Delta v^\sigma\varphi_t \mathrm{d}t\\ =&\sigma\int_0^\infty \int_{\mathbb{R}} v^\sigma\varphi_t\eta''(v_t^\sigma)|\partial_x v_t^\sigma|^2.\end{aligned}$$

Par la convexité de η, si l'on suppose que la fonction test φ est positive, le second membre de cette égalité est positif et la troisième intégrale du membre de gauche tend vers 0. Par conséquent, en prenant la limite $\sigma \to 0$, on obtient l'inégalité

$$\int_{\mathbb{R}} \eta(v_0)\varphi_0 + \int_0^t \int_{\mathbb{R}} \eta(v_t)\partial_t\varphi_t + \psi(v_t)\partial_x\varphi_t \mathrm{d}t \geq 0. \tag{4.5}$$

Cette inégalité, qui traduit l'évolution irréversible de la solution de (4.2) caractérise la limite v de v^σ parmi toutes les solutions de (4.2). On a donc trouvé la bonne notion de solution :

Définition 4.1. *On appelle solution entropique de l'équation* (4.2) *une fonction v de l'espace $\mathbb{L}^\infty((0,\infty), \mathbb{L}^1((1+x^2)^{-1}))$ satisfaisant l'inégalité* (4.5) *pour toute fonction convexe η, et toute fonction positive régulière à support compact φ.*

4.1.2 Lois de conservation scalaires fractionnaires

Il est naturel de s'intéresser au comportement de l'équation obtenue en ajoutant à l'équation (4.2) non pas un terme de diffusion classique comme dans (4.4), mais un terme de diffusion fractionnaire obtenu à partir d'une puissance non entière du Laplacien $-(-\Delta)^{\frac{\alpha}{2}}$, avec α dans l'intervalle $(0,2)$. Pour des exemples physiques faisant intervenir ce type d'opérateur de diffusion, voir [72]. L'opérateur $(-\Delta)^{\frac{\alpha}{2}}$ peut être défini sur un espace de fonctions suffisamment régulières en passant par la transformée de Fourier :

$$\widehat{(-\Delta)^{\frac{\alpha}{2}}\varphi}(\xi) = |\xi|^\alpha\hat{\varphi}(\xi).$$

On voit sur cette expression que cet opérateur est un opérateur de dérivation d'ordre α. Il peut se récrire sous une forme intégrale :

$$-(-\Delta)^{\frac{\alpha}{2}}\varphi(x) = c_\alpha \int_{\mathbb{R}} \frac{\varphi(x+y) - \varphi(x) - y\varphi'(x)\mathbf{1}_{|y|\leq r}}{|y|^{1+\alpha}}\mathrm{d}y,$$

où la constante c_α est une constante dépendant de la normalisation choisie dans la définition de la transformée de Fourier. Il est à noter que le noyau $\mathbf{1}_{|y|\leq r}$ sert à faire converger l'intégrale qui sinon serait singulière, mais que la valeur de r ne joue pas sur la définition de l'opérateur, la mesure $\frac{\mathrm{d}z}{|z|^{1+\alpha}}$ étant symétrique par rapport à l'origine.

Ce type d'opérateur est un cas particulier d'*opérateur de Lévy*. Un tel opérateur est de la forme

$$\mathcal{L}\varphi(x) = \int_{\mathbb{R}} \left(\varphi(x+y) - \varphi(x) - y\varphi'(x)\mathbf{1}_{|y|\leq 1}\right)\mathrm{d}\nu(y),$$

où ν, appelée *mesure de Lévy* de $\mathcal{L}$ est une mesure intégrant la fonction $1 \wedge y^2$. Dans le cas du Laplacien fractionnaire, la mesure de Lévy est $c_\alpha\frac{\mathrm{d}y}{|y|^{1+\alpha}}$.

On peut alors s'intéresser à l'équation

$$\partial_t v_t(x) + \partial_x(A(v_t(x))) = -\sigma(-\Delta)^{\frac{\alpha}{2}}v_t(x), \quad (x,t) \in \mathbb{R}\times(0,\infty). \tag{4.6}$$

La valeur du paramètre α joue un rôle crucial sur le comportement des solutions de (4.6). En effet, on a vu au paragraphe 4.1.1 que la loi de conservation (4.2) non visqueuse peut créer des

chocs, alors que l'ajout d'un terme de diffusion régularise la solution. En conséquence, les deux termes de diffusion et de transport jouent des rôles antagonistes, et le comportement des solutions de l'équation (4.6) va se déduire de la force relative de chaque terme. Le terme de diffusion est d'ordre α, alors que le terme de transport est d'ordre 1, par conséquent, si α est plus grand que 1 le comportement global de la solution de (4.6) sera diffusif et les solutions seront régulières, alors que si α est plus petit que 1, le terme de transport domine et l'on peut observer des chocs. Cette apparition de chocs dépend en réalité de la taille des conditions initiales, voir par exemple [5].

On pourrait également remplacer dans (4.6) le Laplacien fractionnaire par un opérateur de Lévy $\mathcal{L}$ quelconque. La distinction entre opérateur suffisamment régularisant ou pas se ferait alors sur l'intégrabilité de la mesure de Lévy au voisinage de 0 : si $\int_{\mathbb{R}} |y| \wedge 1 \mathrm{d}\nu(y) = \infty$, la diffusion est dominante par rapport au terme de transport. Si $\int_{\mathbb{R}} |y| \wedge 1 \mathrm{d}\nu(y) < \infty$ le terme de transport domine.

Au vu de ces considérations, il n'est pas étonnant que l'analyse mathématiques du cas $\alpha > 1$ soit plus simple que celle du cas $\alpha \leq 1$. Dans le premier cas on a en fait existence et unicité d'une solution faible pour l'équation (4.6), définie comme une fonction satisfaisant l'égalité

$$\int_{\mathbb{R}} v_0 \varphi_0 + \int_0^\infty \int_{\mathbb{R}} A(v_t) \partial_x \varphi_t \mathrm{d}t + \int_0^\infty \int_{\mathbb{R}} (-\Delta)^{\frac{\alpha}{2}} v_t(x) \varphi_t(x) \mathrm{d}t = 0.$$

pour toute fonction test régulière φ.

Dans le cas contraire, le terme de transport domine l'équation, et il est nécessaire de passer par une généralisation de la notion de solution entropique, définie par Alibaud dans [3].

Définition 4.2. *Une fonction v de $\mathbb{L}^\infty$ est dite solution entropique de* (4.6) *si elle vérifie, pour tout réel positif r, pour toute fonction convexe η et toute fonction positive régulière à support compact φ, l'inégalité*

$$\begin{aligned}
&\int_{\mathbb{R}} \eta(v_0(x)) \varphi_0(x) + \int_0^\infty \int_{\mathbb{R}} \eta(v_t(x)) \partial_t \varphi_t(x) + \psi(v_t(x)) \partial_x \varphi_t(x) \mathrm{d}x \mathrm{d}t \\
&+ c_\alpha \int_0^\infty \int_{\mathbb{R}} \int_{\{|y|>r\}} \eta'(v_t(x)) (v_t(x+y) - v_t(x)) \varphi_t(x) \frac{\mathrm{d}y \mathrm{d}x \mathrm{d}t}{|y|^{1+\alpha}} \\
&+ c_\alpha \int_0^\infty \int_{\mathbb{R}} \int_{\{|y|\leq r\}} \eta(v_t(x)) (\varphi_t(x+y) - \varphi_t(x) - y \partial_x \varphi_t(x)) \frac{\mathrm{d}y \mathrm{d}x \mathrm{d}t}{|y|^{1+\alpha}} \geq 0,
\end{aligned}$$

où ψ est une primitive de $\eta' A'$.

On peut remarquer que cette formulation est composée, outre des termes déjà présents dans la formulation entropique de l'équation (4.2), de deux intégrales, l'une portant sur les grandes valeurs du paramètre y, l'autre sur les petites valeurs. Grâce à la convexité de la fonction η utilisée dans la formulation entropique, on voit que si la formulation est vérifiée pour un paramètre r donné, elle est aussi vérifiée pour tous les r' tels que $r' \geq r$. Par conséquent, la formulation entropique doit être comprise dans la limite $r \to 0$.

Cette notion de solution entropique est effectivement nécessaire pour assurer l'unicité de la solution. En effet dans [4] une solution faible qui n'est pas solution entropique est exhibée.

4.2 Interprétation probabiliste des lois de conservation scalaires

4.2.1 Interprétation probabiliste par dérivation en espace

La perturbation (4.4) de la loi de conservation (4.2) décrite dans la section précédente étant une équation parabolique non linéaire, on peut se demander si une interprétation probabiliste de cette équation existe, ainsi que pour les équations (4.2) et (4.6).

Pour trouver cette interprétation probabiliste, il faut en fait considérer non pas les équations (4.2), (4.4) ou (4.6) elles-mêmes, mais en fait l'équation satisfaite par la dérivée $u = \partial_x v$ de leurs solutions. Les trois équations se récrivent alors

$$\partial_t u_t(x) + \partial_x(A'(H * u_t(x))u_t(x))) = 0, \tag{4.7}$$

$$\partial_t u_t(x) + \partial_x(A'(H * u_t(x))u_t(x))) = \frac{\sigma^2}{2}\Delta u_t(x) \tag{4.8}$$

et

$$\partial_t u_t(x) + \partial_x(A'(H * u_t(x))u_t(x))) = -\frac{\sigma^\alpha}{2}(-\Delta)^{\frac{\alpha}{2}} u_t(x). \tag{4.9}$$

Dans ces trois équations, H désigne la fonction de Heaviside $H = \mathbf{1}_{[0,\infty)}$, de sorte que $H * u$ est la primitive de u.

Formellement, si u_0 est une mesure de probabilité, c'est-à-dire si v_0 est une fonction croissante avec des limites 0 et 1 en $\pm\infty$, l'équation (4.8) correspond à l'équation de Fokker-Plank associée à une solution de l'équation

$$\begin{cases} \mathrm{d}X_t &= A'(H * u_t(X_t))\mathrm{d}t + \sigma \mathrm{d}W_t, \\ X_0 & \text{ de loi } u_0 \end{cases} \tag{4.10}$$

où u_t est la loi à l'instant t de X_t. On a donc une équation différentielle stochastique non linéaire au sens de McKean, avec une dérive dépendant de la fonction de répartition de la solution. On reconnaît une perturbation de l'équation des caractéristiques (4.3) par un mouvement Brownien. Si l'on remplace la loi exacte u_t par la loi empirique d'un système de particules $(X_t^i)_{i\in\{1,\dots,N\}}$, le système d'équations satisfait par le système de particules est

$$\begin{cases} \mathrm{d}X_t^i &= A'\left(\frac{1}{N}\sum_{j=1}^N \mathbf{1}_{X_t^j \le X_t^i}\right)\mathrm{d}t + \sigma \mathrm{d}W_t^i \\ (X_0^i)_{i=1,\dots,N} & \textit{i.i.d.} \text{ de loi } u_0 \end{cases},$$

les particules interagissent donc à travers leur fonction de répartition empirique.

Cette interprétation se généralise immédiatement au cas où u_0 est une mesure positive quelconque, c'est-à-dire si v_0 est une fonction croissante bornée, quitte à renormaliser les coefficients de l'équation. On peut toutefois généraliser cette approche au cas où la condition initiale est à variations bornées et non constante, en suivant la méthode de [34]. En effet, la dérivée u_0 de la condition initiale sera alors la différence de deux mesures finies, et on pourra supposer, quitte à renormaliser, que la variation totale de cette mesure est une mesure de probabilité $|u_0|$. La dérivée de Radon-Nikodym $\gamma = \frac{\mathrm{d}u_0}{\mathrm{d}|u_0|}$ est u_0-presque partout à valeurs dans $\{-1, 1\}$. Cette fonction γ correspond alors à un signe attaché à chaque particule. Dans le système de particules, l'approximation naturelle de u_t est alors $\frac{1}{N}\sum_{i=1}^N \mathbf{1}_{X_t^i \le x}\gamma(X_0^i)$, qui est une fonction de répartition empirique signée du système de particules. Le système d'équations différentielles stochastiques décrivant l'évolution du système de particules est alors

$$\begin{cases} \mathrm{d}X_t^i = A'\left(\frac{1}{N}\sum_{j=1}^N \mathbf{1}_{X_t^j \le X_t^i}\gamma(X_t^j)\right)\mathrm{d}t + \sigma \mathrm{d}W_t^i \\ (X_0^i)_{i=1,\dots,N} \quad \textit{i.i.d.} \text{ de loi } |u_0| \end{cases}. \tag{4.11}$$

Ce système de particules à été étudié dans [14, 17, 18, 36, 63]. Le système de particules vérifie un résultat de propagation du chaos vers une solution de l'équation différentielle stochastique non linéaire (4.10), et la loi de cette solution vérifie l'équation (4.4). De plus, on peut estimer l'erreur dans l'approximation du processus non linéaire par un système de particules discrétisé en

temps. L'erreur est de l'ordre de $\frac{1}{\sqrt{N}} + \Delta t$, où N est le nombre de particules et Δt est le pas de discrétisation.

On peut signaler également que le comportement en temps long de ce type de système de particules a été étudié, dans [38] et indépendamment dans [56], avec des méthodes différentes.

Quand le paramètre σ tend vers 0, la solution entropique de l'équation (4.4) converge vers la solution de l'équation (4.2), il serait donc naturel que la solution probabiliste converge vers la solution de l'équation non visqueuse dans la limite $\sigma \to 0$.

La dynamique limite du système de particules (4.11) quand σ tend vers 0 à nombre de particules N fixé est la dynamique des particules collantes, voir par exemple [19, 35]. Dans cette dynamique, on tire des particules (X_0^i) selon la loi $|u_0|$, chacune étant initialement dotée d'une masse 1 ou -1 donnée par $\gamma(X_0^i)$. Les particules sont ensuite lancées initialement à la vitesse $A'\left(\frac{1}{N}\sum_{j=1}^{N}\gamma(X_0^j)\mathbf{1}_{X_0^j \leq X_0^i}\right)$, ce qui n'est autre que la dynamique (4.11) où l'on a posé $\sigma = 0$. Quand deux particules se rencontrent, elles se retrouvent collées en une seule particule dont la masse est la somme des masses des particules, et dont la vitesse est telle que le moment cinétique est conservé.

Toutefois on peut aussi penser à faire tendre le coefficient de diffusion vers 0 en même temps que le nombre de particules tend vers l'infini. Dans ce cas, il est nécessaire d'adapter la dynamique pour pouvoir prouver la convergence de la fonction de répartition empirique vers la solution de (4.2). En effet, si deux particules de signes opposés se croisent, leur contribution dans la formulation entropique approchée vérifiée par le système de particules porte le mauvais signe, ce qui fait que l'on ne retrouve pas la formulation entropique en passant à la limite. Au niveau de la solution, un croisement de deux particules de signes opposés se traduirait par une augmentation de la variation totale, ce qui est en opposition avec l'irréversibilité de l'équation (4.4). La solution est alors de tuer tout couple de particules de signes opposés entrant en contact, leur contribution au moment où elles se touchent étant de toutes façons nulle. L'existence de ce système de particules se montre à partir du théorème de Girsanov appliqué à un mouvement Brownien réfléchi sur les bords du simplexe. On a alors également un résultat de convergence de la fonction de répartition empirique signée vers la solution entropique de (4.2), [36].

4.2.2 Processus de Lévy

L'interprétation probabiliste de la loi de conservation à diffusion anormale (4.6) passe par l'interprétation probabiliste de l'opérateur $-(-\Delta)^{\frac{\alpha}{2}}$, qui est un cas particulier d'opérateur de Lévy. Ces opérateurs sont les générateurs infinitésimaux des *processus de Lévy*.

Les processus de Lévy sont par définition des processus à accroissements indépendants et stationnaires. Il s'agit donc d'une généralisation du mouvement Brownien, puisqu'on ne présuppose plus de continuité sur les trajectoires du processus, ce qui amène naturellement à l'apparition de sauts. On peut montrer que la loi d'un processus de Lévy est entièrement caractérisée par la loi de sa position à l'instant 1. On s'intéresse donc uniquement à la fonction caractéristique de la position du processus à l'instant 1. Selon la formule de Lévy-Khintchine, cette fonction caractéristique s'écrit $\varphi_1^{\mathcal{L}}(u) = e^{\Psi(u)}$, où $\Psi(u)$ est de la forme

$$\Psi(u) = i\alpha u - \sigma^2 u^2/2 + \int_{\mathbb{R}} e^{iuz} - 1 - iuz\mathbf{1}_{|z|\leq 1}\nu(\mathrm{d}z),$$

où ν est une mesure intégrant la fonction $1 \wedge z^2$. Le noyau $\mathbf{1}_{|z|\leq 1}$ peut être remplacé par toute fonction équivalente à 1 en zéro et convergeant suffisamment vite vers zéro en l'infini. Dans cette expression, le terme $i\alpha u$ correspond à une dérive α, et le terme $-\sigma^2 u^2/2$ correspond à une diffusion Brownienne de coefficient σ. Le terme intégral correspond à une dynamique de sauts. La mesure ν décrit la distribution des sauts par unité de temps, les différents sauts étant distribués selon une mesure de Poisson d'intensité $\nu(\mathrm{d}u)\mathrm{d}t$. La condition d'intégrabilité de $1 \wedge z^2$ pour ν traduit le fait

que les grands sauts doivent être en nombre fini, tandis que les petits sauts, s'ils peuvent être en nombre infini, doivent être suffisamment petits pour que leur somme correctement recentrée converge dans $\mathbb{L}^2(\Omega)$.

L'interprétation probabiliste de la loi de conservation fractionnaire (4.6) est alors

$$\mathrm{d}X_t = A'(H * u_t(X_t))\mathrm{d}t + \sigma \mathrm{d}L_t, \tag{4.12}$$

où le processus $(L_t)_{t\geq 0}$ a pour générateur $-(-\Delta)^{\frac{\alpha}{2}}$.

A priori on n'a pas de résultat d'existence ou d'unicité pour cette équation différentielle stochastique. On n'a pas non plus existence du système de particules associé, car le processus dirigeant l'équation est un processus de Lévy dont la mesure est singulière. Dans [39], le système de particules analogue de (4.11) est étudié dans le cas $\alpha > 1$, avec une condition initiale monotone, c'est-à-dire avec $\gamma \equiv 1$. Dans ce cas particulier, on peut prouver l'existence du système de particules, ainsi que la propagation du chaos vers une solution de l'équation non linéaire (4.12). On a également convergence de la fonction de répartition empirique vers la solution de (4.6).

Une manière de contourner ce problème de non existence est de passer en temps discret, en utilisant un schéma d'Euler. On obtient alors le schéma suivant

$$\bar{X}^i_{(k+1)h_N} = A'\left(\frac{1}{N}\sum_{j=1}^{N} \mathbf{1}_{\bar{X}^j_{kh_N} \leq \bar{X}^i_{kh_N}} \gamma(X^j_{kh_N})\right) h_N + \sigma(L^i_{(k+1)h_N} - L^i_{kh_N}),$$

où h_N est le pas de temps, que l'on va faire diminuer au fur et à mesure que le nombre de particules augmente, et où les $(L^i_t)_{t\geq 0}$ sont des processus indépendants de générateur $-(-\Delta)^{\frac{\alpha}{2}}$.

Un autre problème se pose au niveau du meurtre des particules. En effet, en temps discret, les particules ne peuvent pas réellement se trouver à la même position. Il faut alors tuer, à chaque pas de discrétisation, les couples de particules de signes opposés se trouvant à une distance inférieure à un certain seuil ε_N.

On montre sous certaines conditions sur la vitesse de convergence des trois paramètres ε_N, h_N et σ_N, que la fonction de répartition empirique signée converge vers la solution de l'équation (4.2) ou (4.6), suivant la limite choisie sur h_N. Les considérations sur la valeur du paramètre α font que les conditions de cette convergence ne sont pas les mêmes si α est plus grand ou plus petit que 1. Pour être précis, on montre les résultats suivants, dans lesquels F^N désigne la fonction de répartition empirique signée du système de particules :

$$F^N_t(x) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{\bar{X}^i_t \leq x} \gamma(X^i_0),$$

le processus $\bar{X}^i_t$ étant obtenu en interpolant linéairement entre les valeurs $\bar{X}^i_{kh_N}$.

Théorème 4.1. *Supposons $0 < \alpha \leq 1$, et soit $\sigma_N \equiv \sigma$ une suite constante. Si les deux suites ε_N et h_N tendent vers 0 et satisfont*

$$N^{-\lambda} \leq 4 \sup_{[-1,1]} |A'| h_N \leq \varepsilon_N, \text{ et } N^{-1/\alpha} \leq N^{-1/\lambda}\varepsilon_N$$

pour un certain réel positif λ, et si, pour $\alpha = 1$, $h_N \leq \varepsilon_N N^{-1/\lambda}$, alors à tout horizon $T > 0$,

$$\lim_{N\to\infty} \int_0^T \mathbb{E}\left\|F^N_t - v_t\right\|_{\mathbb{L}^1\left(\frac{\mathrm{d}x}{1+x^2}\right)} \mathrm{d}t = 0,$$

où v_t est la solution entropique de (4.6).

Théorème 4.2. *Si ε_N, h_N et σ_N tendent toutes trois vers 0 et satisfont*

$$N^{-\lambda} \leq 4 \sup_{[-1,1]} |A'| h_N \leq \varepsilon_N$$

pour un certain $\lambda > 0$*, et si, quand* $\alpha > 1$*,* $\sigma_N \leq \varepsilon_N^{1-\frac{1}{\alpha}} N^{-\frac{1}{\lambda}}$*, alors, pour tout* $T > 0$*,*

$$\lim_{N\to\infty} \int_0^T \mathbb{E} \left\| F_t^N - v_t \right\|_{\mathbb{L}^1\left(\frac{\mathrm{d}x}{1+x^2}\right)} \mathrm{d}t = 0,$$

où v_t *est la solution entropique de* (4.2)*.*

Ici, une hypothèse supplémentaire est nécessaire quand $\alpha > 1$, car dans ce cas le terme dominant le processus est la diffusion, alors que le terme dominant l'équation limite est le terme de transport. Il faut donc s'assurer que la diffusion ne perturbe pas trop la dynamique de transport. Pour $\alpha \leq 1$, en revanche le terme de transport est déjà dominant.

Théorème 4.3. *Supposons* $1 < \alpha \leq 2$*. Si* $\sigma_N \equiv \sigma$ *est constant et que* ε_N *et* h_N *tendent vers 0, alors pour tout* $T > 0$*,*

$$\lim_{N\to\infty} \int_0^T \mathbb{E} \left\| F_t^N - v_t \right\|_{\mathbb{L}^1\left(\frac{\mathrm{d}x}{1+x^2}\right)} \mathrm{d}t = 0,$$

où v_t *et la solution faible de* (4.6)*.*

Partie II

Calculs d'énergies libres en dynamique moléculaire

5

Existence, unicité et convergence d'une approximation particulaire pour le processus de la Force Biaisante Adaptative

Résumé : Nous étudions une méthode de calcul d'énergie libre introduite dans [25, 32], qui est basée sur le comportement en temps long d'une équation différentielle stochastique non-linéaire. Le terme non-linéaire est une espérance conditionnelle par rapport à une des coordonnées de la solution. Dans [48], la convergence en temps long d'une éventuelle solution est prouvée, sous quelques hypothèses de régularité.

Dans le présent article, nous montrons que cette équation a effectivement une unique solution, sous quelques conditions, et nous étudions une approximation particulaire basée sur un estimateur de l'espérance conditionnelle de type Nadaraya-Watson. Ce système de particules converge vers la solution de l'équation non-linéaire quand le nombre de particules tend vers l'infini, et quand le noyau de Nadaraya-Watson converge vers une masse de Dirac.

Nous donnons également une vitesse pour cette convergence, qui est illustrée par quelques simulations numériques dans un cas simple.

Mots-Clés : Calculs d'énergies libres, équations différentielles stochastiques, équation de Fokker-Planck, systèmes de particules en interaction, espérance conditionnelle, estimateur de Nadaraya-Watson

Abstract : We study a free energy computation procedure, introduced in [25, 32], which relies on the long-time behavior of a nonlinear stochastic differential equation. This nonlinearity comes from a conditional expectation computed with respect to one coordinate of the solution. The long-time convergence of the solutions to this equation has been proved in [48], under some existence and regularity assumptions.

In this paper, we prove existence and uniqueness under suitable conditions for the nonlinear equation, and we study a particle approximation technique based on a Nadaraya-Watson estimator of the conditional expectation. The particle system converges to the solution of the nonlinear equation if the number of particles goes to infinity and then the kernel used in the Nadaraya-Watson approximation tends to a Dirac mass.

We derive a rate for this convergence, and illustrate it by numerical examples on a toy model.

Keywords : Free energy calculations, nonlinear stochastic differential equation, Fokker-Planck equation, interacting particle system, conditional expectation, Nadaraya-Watson estimator

Introduction

Free energy computations are an important problem in the field of molecular simulation (see [24]). The difficulty of those computations lies in the fact that most dynamics in molecular simulations are highly metastable : many free energy barriers prevent a good sampling. We study here the adaptive biasing force (ABF) method, which was introduced in [25, 32] to get rid of those metastabilities.

The typical problems one can think about are the study of a structural angle in the conformation of a protein, or the measure of the evolution of a chemical reaction. Mathematically, each configuration of the system is modelized by an element of a high-dimensional state space $\mathcal{D}$, typically an open subset of $\mathbb{R}^d$, which is endowed with a probability measure, called the canonical measure. This measure is given by $(\int_{\mathcal{D}} e^{-\beta V(x)}\mathrm{d}x)^{-1}e^{-\beta V(x)}\mathrm{d}x$, where V denotes the potential energy undergone by the physical system, and β is proportional to the inverse of the temperature of the system.

For some x in the state space, one is interested in a particular quantity, denoted by $\xi(x)$, ξ being assumed to be a smooth function from $\mathcal{D}$ to the one-dimensional torus $\mathbb{T}$. The quantity $\xi(x)$ has to be understood as a coarse-grained information on the system, which is the relevant information for the practitioner. In the examples above, $\xi(x)$ would be a structural angle in a protein with conformation x, or a number measuring the evolution of a chemical system in state x.

We call *free energy* the effective energy associated to the quantity $\xi(x)$, that is, the function $A(z)$ such that $e^{-\beta A(z)}\mathrm{d}z$ is the image measure of the canonical measure by the function ξ. Our objective is to compute numerically the function A. When $\mathcal{D} = \mathbb{R}^d$, a naive method to do so is to simulate, for a given random variable X_0 and an independent $\mathbb{R}^d-$valued Brownian motion W, the process defined by the (overdamped) Langevin dynamics

$$\mathrm{d}X_t = -\nabla V(X_t)\mathrm{d}t + \sqrt{2\beta^{-1}}\mathrm{d}W_t, \tag{5.1}$$

which, under some regularity assumptions on the potential, is ergodic and admits the canonical measure as unique invariant measure. This approach appears to be untractable in practice, since the convergence to equilibrium is very slow, due to multiple metastabilities appearing in most problems : typically, a molecule moves microscopically within times of order 10^{-15} seconds, while the typical time scale of the macroscopic moves is of order 10^{-9} seconds.

The idea of the ABF method is to prevent the process X_t from staying in metastable states by introducing a biasing force which repel X_t from the states where it stayed for too long a time. To do this, we use the following representation of A, that can be deduced from the co-area formula (see [48]) :

$$A'(z) = \mathbb{E}\left[F(X)|\xi(X) = z\right], \tag{5.2}$$

where X is a random variable distributed according to the canonical measure, and F is the function defined by

$$F(x) = \frac{\nabla\xi\cdot\nabla V}{|\nabla\xi|^2} - \frac{1}{\beta}\mathrm{div}\left(\frac{\nabla\xi}{|\nabla\xi|^2}\right). \tag{5.3}$$

The function A' is called the *mean force*. Actually, (5.2) also holds when X is distributed according to the measure

$$\left(\int_{\mathcal{D}} e^{-\beta(V(x)+W\circ\xi(x))}\mathrm{d}x\right)^{-1} e^{-\beta(V(x)+W\circ\xi(x))}\mathrm{d}x,$$

which is the canonical measure associated with the biased potential $V + W\circ\xi$ where W is any smooth function.

Equation (5.2) leads us to consider the following dynamics, which should get rid of metastabilities for a well chosen ξ since it "flattens" the energy landscape in the ξ direction (see [48] and Lemma 5.1.2 below for more precise statements) :

$$\begin{cases} \mathrm{d}X_t &= -\nabla\left(V - A_t \circ \xi - \beta^{-1}\ln(|\nabla\xi|^{-2})\right)(X_t)|\nabla\xi|^{-2}(X_t)\mathrm{d}t + \sqrt{2\beta^{-1}}|\nabla\xi|^{-1}(X_t)\mathrm{d}W_t, \\ A'_t(z) &= \mathbb{E}\left[F(X_t)|\xi(X_t) = z\right]. \end{cases} \tag{5.4}$$

The second equality in (5.4) shows that if X_t is distributed according to the canonical measure associated with the potential $V - A \circ \xi$, then the biasing force A'_t is actually the derivative A' of the free energy, and the first equation in (5.4) consists in a Langevin dynamics associated to the potential $V - A \circ \xi$. Consequently, the dynamics (5.4) admits a stationary point : $A'_t = A'$ and $\mathrm{Law}(X_t) = (\int_{\mathcal{D}} e^{-\beta(V-A\circ\xi)}\mathrm{d}x)^{-1}e^{-\beta(V-A\circ\xi)}\mathrm{d}x$. The diffusion term $|\nabla\xi|^{-1}(X_t)$ in (5.4) (and the associated modifications of the drift term) is required to obtain natural longtime convergence results, but a constant diffusion term can also be used, see [48] for more details.

If we actually have convergence to this stationary state, we have a method, that should be efficient (*i.e.* that should not see the metastabilities), to sample the canonical measure up to a known perturbation $e^{A\circ\xi}$. This algorithm has thus two applications : it allows the computation of the free energy A, and it can be used as an adaptative importance sampling method for the canonical measure.

The long time behavior of Equation (5.4) has been studied in [48], where it has been proven that for a sufficiently regular solution, one has, in some sense, an exponential convergence to the stationary state, with a rate that is better (for a well chosen ξ) than the rate of convergence to equilibrium for (5.1).

The practical difficulty in simulating (5.4) is to compute the conditional expectation, which is a highly nonlinear term. Stochastic differential equations involving conditional expectations have already been studied, in a case where the conditional expectation is computed with respect to a random initial condition (see [66, 69]) or where the variable whose conditional expectation is computed is fixed (see [28]). Our situation is much more complex since both the conditioning and the conditioned variables change with time and are affected by the previous conditional expectations.

The same difficulty arises in Lagrangian stochastic models which are commonly used in the simulation of turbulent flows (see [16]). The main difference between the system studied in [16] and (5.4) is that the authors considers a Langevin dynamics with noise only on the velocity. The lack of ellipticity then leads to additional difficulties. In our setting we are able to derive a quantitative error estimate for the particle discretization while this seems more difficult for Langevin dynamics.

In this paper, we prove that existence and uniqueness hold for Equation (5.4) under suitable conditions, and we study an approximation of X_t by an interacting particle system (see Theorems 5.1.3 and 5.1.4 below).

The paper is organized as follows. In Section 5.1 we state our main results.

Section 5.2 is devoted to some uniqueness and regularity results. More precisely, we prove that the time marginals of a solution to Equation (5.4) satisfy some partial differential equation. Then, under an integrability condition on the initial condition, we prove uniqueness for the solutions to this equation, so that the nonlinear term in (5.4) is reduced to a bounded drift coefficient. We thus prove pathwise uniqueness and uniqueness in distribution for the solutions of (5.4).

Section 5.3 is devoted to existence results. More precisely, we introduce a regularization of the dynamics (5.4) involving two parameters α and ε, which is another nonlinear stochastic differential equation whose nonlinearity is less singular. We prove that strong existence, pathwise uniqueness and uniqueness in distribution hold for this equation and then we show that the solutions to this stochastic differential equation converge to some process which satisfies (5.4) in the limit $(\alpha, \varepsilon) \to (0, 0)$, yielding strong existence. We also prove that this convergence holds with rate $\mathcal{O}(\alpha + \sqrt{\varepsilon})$.

In Section 5.4 we introduce an interacting particle system to approximate the regularized dynamics, and we prove a propagation-of-chaos result for this particle system. We also derive a rate of convergence for this propagation of chaos.

In Section 5.5, we illustrate the efficiency of the particle approximation of the ABF method and the rate of those convergences with some numerical examples in small dimension.

Notation

We denote by $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ the one dimensional torus, and for $x \in \mathbb{R}$, we denote by $\{x\}$ the fractional part of x, that can be seen as a projection of x on $\mathbb{T}$. In the following, we will work in two different domains $\mathcal{D} : \mathbb{T} \times \mathbb{R}^{d-1}$ or $\mathbb{T}^d$. The case $\mathcal{D} = \mathbb{T} \times \mathbb{R}^{d-1}$ will be called the non compact case, and the case $\mathcal{D} = \mathbb{T}^d$ will be called the compact case. For $x \in \mathbb{R}^d$, depending on the case considered, we will also denote by $\{x\}$ the element of $\mathbb{T} \times \mathbb{R}^{d-1}$ (resp. $\mathbb{T}^d$) defined by $\{x\} = (\{x^1\}, x^2, \dots, x^d)$ (resp. $\{x\} = (\{x^1\}, \dots, \{x^d\})$).

In the following, we will call "function defined on $\mathbb{T}$" (resp. on $\mathbb{T} \times \mathbb{R}^{d-1}$, resp. on $\mathbb{T}^d$), a $\mathbb{Z}-$periodical (resp. $\mathbb{Z}-$periodical in the first coordinate, resp. $\mathbb{Z}^d$-periodical) function defined on $\mathbb{R}$ (resp. on $\mathbb{R}^d$). Integrals on $\mathbb{T}$, $\mathbb{T} \times \mathbb{R}^{d-1}$ or $\mathbb{T}^d$ mean integrals on $[0,1)$, $[0,1) \times \mathbb{R}^{d-1}$ or $[0,1)^d$.

We denote by $\mathbb{L}^2(\mathbb{T}^d)$ the space of functions on $\mathbb{T}^d$ whose square is integrable on $\mathbb{T}^d$, and by $\mathbb{H}^1(\mathbb{T}^d)$ the space of functions in $\mathbb{L}^2(\mathbb{T}^d)$ whose weak gradient is square integrable on $\mathbb{T}^d$. We use similar notations on $\mathbb{T} \times \mathbb{R}^{d-1}$ and $\mathbb{T}$.

For two functions f and g defined on $\mathbb{T} \times \mathbb{R}^{d-1}$ or $\mathbb{T}^d$, we denote $f * g$ the convolution *with respect to the first coordinate*, that is,

$$f * g(x) = \int_{\mathbb{T}} f(x^1 - y^1, x^{2\dots d}) g(y^1, x^{2\dots d}) \mathrm{d}y^1.$$

If f is defined on $\mathbb{T}$, we also use the notation $f * g$ to denote

$$f * g(x) = \int_{\mathbb{T}} f(x^1 - y^1) g(y^1, x^{2\dots d}) \mathrm{d}y^1.$$

When f and g are defined on $\mathcal{D} = \mathbb{T} \times \mathbb{R}^{d-1}$ or $\mathbb{T}^d$, the convolution in *all* the coordinates is denoted $f \star g$:

$$f \star g(x) = \int_{\mathcal{D}} f(x^1 - y^1, x^{2\dots d} - y^{2\dots d}) g(y^1, y^{2\dots d}) \mathrm{d}y^1 \mathrm{d}y^{2\dots d}.$$

In the following, we call "probability measure on $\mathbb{T}$" (resp. on $\mathbb{T} \times \mathbb{R}^{d-1}$, $\mathbb{T}^d$) a nonnegative $\mathbb{Z}$-periodical (resp. $\mathbb{Z}$-periodical with respect to the first coordinate, $\mathbb{Z}^d$-periodical) measure μ such that $\mu([0,1)) = 1$ (resp. $\mu([0,1) \times \mathbb{R}^{d-1}) = 1$, $\mu([0,1)^d) = 1$).

When $\{X\}$ is a random variable taking values in $\mathbb{T}$ (resp. in $\mathbb{T} \times \mathbb{R}^{d-1}$, $\mathbb{T}^d$), we call "distribution of $\{X\}$" or "law of $\{X\}$" the probability measure μ on $\mathbb{T}$ (resp. on $\mathbb{T} \times \mathbb{R}^{d-1}$, $\mathbb{T}^d$) such that

$$\mathbb{E}[f(\{X\})] = \int f(x) \mu(\mathrm{d}x).$$

For a given probability measure μ on $\mathbb{T} \times \mathbb{R}^{d-1}$ (resp. a probability density u) and a given bounded function g, we denote μ^g (resp. $u^g(x^1)\mathrm{d}x^1$) the marginal on $\mathbb{T}$ of the measure $g.\mu$ (resp. $g(x)u(x)\mathrm{d}x$). Namely :

$$\mu^g(A) = \int_{A \times \mathbb{R}^{d-1}} g \mathrm{d}\mu$$

and

$$u^g(x^1) = \int_{\mathbb{R}^{d-1}} g(x^1, x^{2\dots d}) u(x^1, x^{2\dots d}) \mathrm{d}x^{2\dots d}.$$

In particular, μ^1 is the first coordinate marginal of μ. When we do not specify the measure in an integral, it is the Lebesgue measure.

We will need the weighted spaces

$$\mathbb{L}^p(w) = \left\{ \psi \in \mathbb{L}^p(\mathbb{T}\times\mathbb{R}^{d-1}) \text{ s.t. } \|\psi\|_{\mathbb{L}^p(w)} \stackrel{\text{def}}{=} \left(\int_{\mathbb{T}\times\mathbb{R}^{d-1}} |\psi|^p w \right)^{1/p} < \infty \right\},$$

for $1 \le p < \infty$, and

$$\mathbb{H}^1(w) = \left\{ \psi \in \mathbb{H}^1(\mathbb{T}\times\mathbb{R}^{d-1}) \text{ s.t. } \|\psi\|_{\mathbb{H}^1(w)} \stackrel{\text{def}}{=} \left(\int_{\mathbb{T}\times\mathbb{R}^{d-1}} \left(|\psi|^2 + |\nabla\psi|^2\right) w \right)^{1/2} < \infty \right\}$$

with $w(x) = (1+|x^{2\dots d}|^2)^\lambda$, for some $\lambda > (d-1)/2$. Notice that w does not depend on the first coordinate x^1, and that there is a positive constant K such that

$$\forall x \in \mathbb{T}\times\mathbb{R}^{d-1},\ |\nabla w(x)| \le 2\lambda(1+|x^{2\dots d}|^2)^{\lambda-1} \sum_{i=2}^{d} |x^i| \le K w(x). \tag{5.5}$$

We will use several times the following statement :

Lemma 5.0.1 *For a bounded function g, and $u \in \mathbb{L}^2(w)$ one has, for some constant K,*

$$\|u^g\|_{\mathbb{L}^2(\mathbb{T})} \le K \|g\|_{\mathbb{L}^\infty(\mathbb{T}\times\mathbb{R}^{d-1})} \|u\|_{\mathbb{L}^2(w)}.$$

If moreover, g has bounded derivatives and $u \in \mathbb{H}^1(w)$, then

$$\|u^g\|_{\mathbb{H}^1(\mathbb{T})} \le K \|g\|_{\mathbb{W}^{1,\infty}(\mathbb{T}\times\mathbb{R}^{d-1})} \|u\|_{\mathbb{H}^1(w)}.$$

The same inequalities hold with the non weighted norms in the right-hand side, for u respectively in $\mathbb{L}^2(\mathbb{T}^d)$ and $\mathbb{H}^1(\mathbb{T}^d)$.

Proof. Recall that we assumed $\lambda > \frac{d-1}{2}$, so that $\frac{1}{w}$ is integrable on $\mathbb{R}^d$: $\int_{\mathbb{R}^d} \frac{1}{w} \mathrm{d}x < \infty$. Consequently, we have the estimation

$$\begin{aligned}
\|u^g\|^2_{\mathbb{L}^2(\mathbb{T})} &= \int_{\mathbb{T}} \left| \int_{\mathbb{R}^{d-1}} gu \right|^2 \\
&\le \|g\|^2_{\mathbb{L}^\infty(\mathbb{T}\times\mathbb{R}^{d-1})} \int_{\mathbb{T}} \left(\int_{\mathbb{R}^{d-1}} |u|^2 w \int_{\mathbb{R}^{d-1}} \frac{1}{w} \right) \\
&\le K \|g\|^2_{\mathbb{L}^\infty(\mathbb{T}\times\mathbb{R}^{d-1})} \|u\|^2_{\mathbb{L}^2(w)}.
\end{aligned}$$

The proof is similar in the space $\mathbb{H}^1(w)$.

In the following, K will denote some positive constant, whose value can change from line to line.

5.1 Assumptions and statement of the main results

In this paper, we consider a particular case of Equation (5.4) to simplify the argumentation : we assume $\beta = 1$ (this can be realized by a change of variable), $\mathcal{D} = \mathbb{T}\times\mathbb{R}^{d-1}$ or $\mathcal{D} = \mathbb{T}^d$. We consider as reaction coordinate the first coordinate function $\xi : \mathcal{D} \to \mathbb{R}$ defined by $\xi(x) = \xi(x^1, x^2, \dots, x^d) = x^1$. This should not change the theoretical results, but will simplify the proofs. The definition (5.3) of F is then reduced to

$$F = \partial_1 V,$$

where V is defined on $\mathbb{T}^d$ or $\mathbb{T} \times \mathbb{R}^{d-1}$.

The two settings $\mathcal{D} = \mathbb{T}^d$ and $\mathcal{D} = \mathbb{T} \times \mathbb{R}^{d-1}$ will be respectively called the *compact* and the *non-compact* case. Our results hold in both settings, and the proofs are mostly identical, with some slight additional difficulties in the non compact case. Thus, in those situations, we only give the proofs in the non-compact case.

With those assumptions, Equation (5.4) rewrites

$$\mathrm{d}X_t = \big(- \nabla V(X_t) + \mathbb{E}\left[\partial_1 V(X_t) | \{X_t^1\}\right] e_1 \big)\mathrm{d}t + \sqrt{2}\mathrm{d}W_t, \tag{5.6}$$

e_1 denoting the first vector in the canonical basis of $\mathbb{R}^d$. We will call solution to Equation (5.6) a process $\{X_t\}$ where X_t satisfies (5.6). The initial condition of (5.6) is a random variable denoted X_0, and is supposed to be independent of the Brownian motion W. We denote by P_0 the law of $\{X_0\}$, which is a probability measure on $\mathcal{D}$.

To ensure the integrability of $\partial_1 V(X_t)$, we make the following assumption :

Assumption i *V is a twice continuously differentiable function, which has bounded first and second order partial derivatives.*

Notice that Assumption i yields boundedness of the drift coefficient in (5.6). In the compact case, assumption i is satisfied as soon as V is a twice differentiable function.

We have to make some assumptions on the initial condition X_0. What is needed to prove our results will depend on whether we consider the compact or the non compact case. In the compact case, we consider the following assumption :

Assumption ii *The probability measure P_0 has a density p_0 lying in $\mathbb{L}^2(\mathbb{T}^d)$ and whose first coordinate marginal p_0^1 is bounded from below by a positive constant. (Notice that p_0^1 is a probability density on $\mathbb{T}$.)*

In the non compact case, we will need a stronger assumption : we have to control the decay of the initial condition at infinity, so we work in the weighted space $\mathbb{L}^2(w)$. We will use, in addition to Assumption ii, the following one :

Assumption iii *The density p_0 of P_0 lies in both $\mathbb{L}^1(w)$ and $\mathbb{L}^2(w)$.*

Notice that Assumptions iii implies that $\{X_0\}$ has finite moments of order less than 2λ, and that Assumption i then yields a control on the corresponding moments of any solution to (5.6), uniformly in $t \in \mathbb{R}$:

Lemma 5.1.1 *Under Assumptions i and iii, on any bounded time interval $[0, T]$, the moments of order less than 2λ of any solution X of (5.6) are bounded :*

$$\sup_{0 \le t \le T} \mathbb{E}[|X_t|^{2\lambda}] < \infty.$$

Proof. This comes from the boundedness of the drift coeficient $b_s(x) = -\nabla V(x) + \mathbb{E}[\partial_1 V(X) | X^1 = x^1]$, which holds in regard of Assumption i. Indeed, we have $\mathbb{E}[|X_t|^{2\lambda}] = \mathbb{E}[|X_0 + \int_0^t b_s(X_s)\mathrm{d}s + \sqrt{2}W_t|^{2\lambda}] \le K\left(\mathbb{E}[|X_0|^{2\lambda}] + t^{2\lambda} + t^{\lambda}\right)$, which is bounded on $[0, T]$.

According to the following fundamental lemma, the solution to (5.6) samples efficiently the coordinate reaction state space $\mathbb{T}$.

Lemma 5.1.2 *Denote by P_t the law of $\{X_t\}$, where X_t is a solution to Equation (5.6). Then, P_t^1 has a density p_t^1, such that p^1 satisfies the heat equation on $\mathbb{T}$ with initial condition p_0^1. Thus, p^1 is uniquely defined on $\mathbb{T} \times [0, \infty)$, and smooth on $\mathbb{T} \times (0, \infty)$.*

Proof (Proof of Lemma 5.1.2). Let f be a smooth function on $\mathbb{T}$. One has, by Itō's formula

$$\partial_t \mathbb{E}\left[f(X_t^1)\right] = -\mathbb{E}\left[f'(X_t^1)\partial_1 V(X_t)\right] + \mathbb{E}\left[f'(X_t^1)\mathbb{E}\left[\partial_1 V(X_t)|\{X_t^1\}\right]\right] + \mathbb{E}\left[f''(X_t^1)\right].$$

But, f being a function on $\mathbb{T}$, $f'(X_t^1)$ only depends on $\{X_t^1\}$, so that the two first terms in the right hand side cancel. Then, it holds that

$$\partial_t \mathbb{E}\left[f(X_t^1)\right] = \mathbb{E}\left[f''(X_t^1)\right],$$

which is exactly the heat equation in the weak sense for $t \mapsto p_t^1$, p_t^1 being the distribution of $\{X_t^1\}$. For uniqueness and regularity of this solution, see [26, Chapter XIV].

Lemma 5.1.2 allows us to rewrite equation (5.6) using the distribution of $\{X_t^1\}$. Indeed, since P_t^1 has a density, the measure given for $A \subset [0,1)$ by $P_t^{\partial_1 V}(A) = \mathbb{E}\left[\partial_1 V(X_t)\mathbf{1}_A(\{X_t^1\})\right]$ also has a density $p_t^{\partial_1 V}$. We can thus write

$$\begin{cases} \mathrm{d}X_t &= \left(-\nabla V(X_t) + \frac{p_t^{\partial_1 V}(X_t^1)}{p_t^1(X_t^1)} e_1\right)\mathrm{d}t + \sqrt{2}\mathrm{d}W_t, \\ P_t &= \text{distribution of } \{X_t\}. \end{cases} \tag{5.7}$$

Moreover, under Assumption ii the density p_t^1 satisfies $0 < \inf_{\mathbb{T}} p_0^1 \leq p_t^1$, uniformly in time, thanks to the maximum principle. This assumption will consequently prevent the denominator in the second term of (5.7) from vanishing.

In view of Equation (5.7), a natural particle approximation of X_t is then obtained using the Nadaraya-Watson estimator of a conditional expectation (see [70]), given, for some parameter η and for a positive integer N, by the system of N stochastic differential equations

$$\mathrm{d}X_{t,n,N}^\eta = \left(-\nabla V(X_{t,n,N}^\eta) + \frac{\sum_{m=1}^N \varphi_\eta(X_{t,n,N}^{\eta,1} - X_{t,m,N}^{\eta,1})\partial_1 V(X_{t,m,N}^\eta)}{\sum_{m=1}^N \varphi_\eta(X_{t,n,N}^{\eta,1} - X_{t,m,N}^{\eta,1})} e_1\right)\mathrm{d}t + \sqrt{2}\mathrm{d}W_t^n,\ 1 \leq n \leq N \tag{5.8}$$

where (W_t^n) is a sequence of independent Brownian motions, and φ_η is a smooth approximation for the Dirac measure at the origin on $\mathbb{T}$. For the initial condition, we work with the following assumption

Assumption iv *The initial condition of Equation* (5.8) *is* $(X_{0,n,N}^\eta)_{0\leq n\leq N} = (X_{0,n})_{0\leq n\leq N}$*, where* $(X_{0,n})_{n\in\mathbb{N}}$ *is a sequence of i.i.d random variables with density* p_0*, and independent of the Brownian motions* $(W_t^n)_{t\geq 0}$.

We also need an assumption on the shape of φ_η. The parameter $\eta = (\alpha, \varepsilon)$ will be chosen in $(0,\infty)^2$, and φ_η will have the form

$$\varphi_\eta(x) = \alpha + \psi_\varepsilon(x), \tag{5.9}$$

where ψ_ε is a sequence of mollifiers on $\mathbb{T}$ as $\varepsilon \to 0$. Namely, assuming $\varepsilon < 1/2$, ψ_ε is a smooth non-negative $\mathbb{Z}$-periodical function, such that $\psi_\varepsilon \equiv 0$ on $[-1/2, 1/2] \setminus [-\varepsilon, \varepsilon]$ and such that

$$\int_{-1/2}^{1/2} \psi_\varepsilon = 1.$$

A simple way to construct such a sequence is to consider a smooth non-negative function ψ defined on $\mathbb{R}$, with support in $[-1,1]$ such that $\int_{\mathbb{R}} \psi = 1$, and then consider the $\mathbb{Z}$-periodization ψ_ε of $\psi_\varepsilon = \frac{1}{\varepsilon}\psi(\frac{\cdot}{\varepsilon})$ (ψ_ε is well defined for $\varepsilon < 1/2$). This example makes the following assumption natural :

Assumption v *The function ψ_ε satisfies*

$$\|\psi_\varepsilon\|_{\mathbb{L}^\infty(\mathbb{T})} \leq \frac{K}{\varepsilon}, \text{ and } \|\psi'_\varepsilon\|_{\mathbb{L}^\infty(\mathbb{T})} \leq \frac{K}{\varepsilon^2}.$$

The reason for adding a positive constant α to the mollifier is to avoid singularities at the denominator in the right-hand side of (5.8). Notice that (5.9) yields strong existence and uniqueness for (5.8), since the drift is globally Lipschitz continuous.

We are going to prove the following two results :

Theorem 5.1.3 *[Existence and uniqueness of the solution] In both the compact and non compact cases, under Assumption i, weak existence holds for Equation* (5.6)*. If P denotes the distribution of a solution, then for all $s > 0$ the time marginals P_s of P admits a density p_s, such that for all $0 < t < T$,*

$$p \in \mathbb{L}^\infty((t,T), \mathbb{L}^2(\mathcal{D})) \bigcap \mathbb{L}^2((t,T), \mathbb{H}^1(\mathcal{D})). \tag{5.10}$$

Moreover, under both Assumptions i and ii for the compact case, and under Assumptions i, ii and iii for the non compact case, strong existence, pathwise uniqueness and uniqueness in distribution also hold, and one can take $t = 0$ in (5.10)*.*

Theorem 5.1.4 *[Particle approximation of the process X_t] Let us consider the processes $X_{t,n,N}$ defined by* (5.8)*. Then, under Assumptions i, ii, iv and v in the compact case, and the additional Assumption iii in the non-compact case, it holds that, for any positive T, and for α and ε small enough,*

$$\mathbb{E}\left[\int_0^T \left\| \frac{\sum_{n=1}^N \partial_1 V(X^\eta_{t,n,N}) \varphi_\eta(. - X^{\eta,1}_{t,n,N})}{\sum_{n=1}^N \varphi_\eta(. - X^{\eta,1}_{t,n,N})} - A'_t \right\|_{\mathbb{L}^\infty(\mathbb{T})} \mathrm{d}t\right] = \mathcal{O}\left(\alpha + \sqrt{\varepsilon} + \frac{1}{\sqrt{N}} e^{\frac{K}{\alpha\varepsilon^2}}\right).$$

Theorem 5.1.3 is a consequence of Theorem 5.2.8 and Corollary 5.3.10 below, and Theorem 5.1.4 is a consequence of Theorems 5.3.11 and 5.4.1 below.

The convergence rate in Theorem 5.1.4 is certainly not optimal. Indeed, it is natural that, for the error to vanish, the number N of particles should go to infinity as ε goes to zero, but the dependency of N on ε which is required for the control of the error in Theorem 5.1.4 to go to zero is certainly pessimistic. This is discussed more precisely in Section 5.5.

5.2 Notion of solution, regularity and uniqueness results

In this section we consider the Fokker-Planck equation associated to the nonlinear stochastic differential equation (5.6) and prove that uniqueness holds for weak solutions of this partial differential equation. From this uniqueness result, the study of Equation (5.6) can be reduced to the study of a linear stochastic differential equation. We can thus prove uniqueness for Equation (5.6).

Let us derive the Fokker-Planck equation associated to Equation (5.6). Let ψ be a twice continuously differentiable function. Applying Itō's formula and taking the expectation, we obtain that the law P_t of a weak solution $\{X_t\}$ to equation (5.6) satisfies

$$\int_{\mathcal{D}} \psi(x) \mathrm{d}P_T(x) = \int_{\mathcal{D}} \psi(x) \mathrm{d}P_0(x) - \int_0^T \int_{\mathcal{D}} \nabla\psi(x) \cdot \nabla V(x) \mathrm{d}P_t(x)\mathrm{d}t + \int_0^T \int_{\mathcal{D}} \Delta\psi(x) \mathrm{d}P_t(x)\mathrm{d}t \tag{5.11}$$

$$+ \int_0^T \int_{\mathcal{D}} \partial_1\psi(x) \left(\frac{p_t^{\partial_1 V}}{p_t^1}(x^1)\right) \mathrm{d}P_t(x)\mathrm{d}t,$$

which is a weak formulation of the following partial differential equation

$$\partial_t P_t = \operatorname{div}\left(P_t \nabla V + \nabla P_t\right) - \partial_1 \left(P_t \frac{p_t^{\partial_1 V}}{p_t^1} \right), \tag{5.12}$$

with initial condition P_0. Using integration by parts, we introduce a stronger definition for solutions to (5.12) which will allow us to prove existence and uniqueness.

Definition 5.2.1 *In the compact case, a function u is said to be a solution to (5.12) if, for any positive T,*

- *u belongs to $\mathbb{L}^\infty((0,T),\mathbb{L}^2(\mathbb{T}^d)) \bigcap \mathbb{L}^2((0,T),\mathbb{H}^1(\mathbb{T}^d))$;*
- *for any function $\psi \in \mathbb{H}^1(\mathbb{T}^d)$, we have :*

$$\partial_t \int_{\mathcal{D}} u_t \psi = -\int_{\mathcal{D}} u_t \nabla V \cdot \nabla \psi - \int_{\mathcal{D}} \nabla u_t \cdot \nabla \psi + \int_{\mathcal{D}} u_t \frac{u_t^{\partial_1 V}}{u_t^1} \partial_1 \psi, \tag{5.13}$$

  in the sense of distributions in time ;
- *$u_0 = p_0$.*

In the non compact case, u is said to be a solution to (5.12), if, for any positive T,

- *u belongs to $\mathbb{L}^\infty((0,T),\mathbb{L}^2(w)) \bigcap \mathbb{L}^2((0,T),\mathbb{H}^1(w))$;*
- *for any $\psi \in \mathbb{H}^1(w)$*

$$\partial_t \int_{\mathcal{D}} u_t \psi w = -\int_{\mathcal{D}} u_t \nabla V \cdot (w\nabla\psi + \psi \nabla w) - \int_{\mathcal{D}} \nabla u_t \cdot (w\nabla\psi + \psi\nabla w) + \int_{\mathcal{D}} u_t \frac{u_t^{\partial_1 V}}{u_t^1} (\partial_1 \psi) w, \tag{5.14}$$

  holds in the sense of distributions in time ;
- *$u_0 = p_0$.*

Notice that (5.13) is a variational formulation of (5.12) in the space $\mathbb{L}^2(\mathbb{T}^d)$ and that (5.14) is a variational formulation of (5.12) in the space $\mathbb{L}^2(w)$.

These conditions make sense. Indeed, in both cases, the conditions on u and ψ are such that the variational formulations (5.13) and (5.14) are well defined (notice that one has $|\nabla w| \leq K w$). Moreover, for the compact case, if u lies in $\mathbb{L}^2((0,T),\mathbb{H}^1(\mathbb{T}^d))$, and satisfies (5.13) then $\partial_t u$ lies in $\mathbb{L}^2((0,T),\mathbb{H}^{-1}(\mathbb{T}^d))$, so that (see [51, page 23]) u lies in $\mathcal{C}([0,T],\mathbb{L}^2(\mathbb{T}^d))$, allowing us to define the value of u at time $t=0$. The same argument holds for the non compact case.

5.2.1 Existence of regular densities for solutions to the nonlinear equation

In this section, we consider a solution X to Equation (5.6) and we denote by P_t the law of $\{X_t\}$. We show that P_t has a density p_t, and that p is a solution to Equation (5.12), in the sense of Definition 5.2.1.

Lemma 5.2.2 *Consider both the compact and the non compact cases. Under Assumption i, for any $t \geq 0$, P_t admits a density p_t with respect to the Lebesgue measure satisfying the following mild representation*

$$p_t = G_t \star P_0 + \int_0^t \nabla G_{t-s} \star (\nabla V p_s) \mathrm{d}s - \int_0^t \partial_1 G_{t-s} \star \left(\frac{p_s^{\partial_1 V}}{p_s^1} p_s \right) \mathrm{d}s, \tag{5.15}$$

where G_t is the density of $\sqrt{2}$ times the Brownian motion on $\mathcal{D}$, namely

$$G_t(x) = \frac{1}{(4\pi t)^{d/2}} \sum_{k \in \mathbb{Z}} e^{-\frac{|x - k e_1|^2}{4t}}$$

for the non-compact case, and

$$G_t(x) = \frac{1}{(4\pi t)^{d/2}} \sum_{k\in\mathbb{Z}^d} e^{-\frac{|x-k|^2}{4t}}$$

for the compact case.

Proof. Let χ be a smooth function with compact support on $\mathbb{T}\times\mathbb{R}^{d-1}$ and $T>0$. Then, for $t\in[0,T]$, the function ψ defined by

$$\psi_s = G_{t-s} \star \chi,$$

is the unique smooth solution to the following problem

$$\begin{cases} \partial_s \psi &= -\Delta\psi \text{ on } (0,t)\times\mathbb{T}\times\mathbb{R}^{d-1}, \\ \psi_t &= \chi \qquad \text{on } \mathbb{T}\times\mathbb{R}^{d-1}. \end{cases} \tag{5.16}$$

Computing $\psi_s(X_s)$ by Itō's formula and using (5.16) we get

$$\begin{aligned} \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \psi_t \mathrm{d}P_t &= \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \psi_0 \mathrm{d}P_0 - \int_0^t \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \Delta\psi_s \mathrm{d}P_s \mathrm{d}s + \int_0^t \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \Delta\psi_s \mathrm{d}P_s \mathrm{d}s \\ &\quad - \int_0^t \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \nabla\psi_s \cdot \nabla V \mathrm{d}P_s \mathrm{d}s + \int_0^t \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \partial_1\psi_s \frac{p_s^{\partial_1 V}}{p_s^1} \mathrm{d}P_s \mathrm{d}s \\ &= \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \psi_0 \mathrm{d}P_0 - \int_0^t \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \nabla\psi_s \cdot \nabla V \mathrm{d}P_s \mathrm{d}s + \int_0^t \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \partial_1\psi_s \frac{p_s^{\partial_1 V}}{p_s^1} \mathrm{d}P_s \mathrm{d}s. \end{aligned}$$

Using the expression of ψ_t and Fubini's Theorem, we have :

$$\begin{aligned} \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \chi \mathrm{d}P_t &= \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \chi (G_t \star P_0) + \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \chi \int_0^t \nabla G_{t-s} \star (P_s \nabla V) \mathrm{d}s \\ &\quad - \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \chi \int_0^t \partial_1 G_{t-s} \star \left(\frac{p_s^{\partial_1 V}}{p_s^1} P_s \right) \mathrm{d}s. \end{aligned}$$

This last equation being true for any smooth function χ with compact support, then P_t is given by the right-hand side of (5.15), which is an integrable function, so that for any positive t, P_t has a density p_t satisfying (5.15).

In regard of the following lemma, p necessarily satisfies some integrability conditions.

Lemma 5.2.3 *In both the compact and the non compact case, under Assumptions i and ii, p lies in $\mathbb{L}^\infty((0,T),\mathbb{L}^2(\mathcal{D}))$ for any $T>0$, and we have $\|p\|_{\mathbb{L}^\infty((0,T),\mathbb{L}^2(\mathcal{D}))} \le C$, where C is some constant only depending on P_0, ∇V and T.*

In the non compact case, under Assumptions i, ii and iii, p lies in $\mathbb{L}^\infty((0,T),\mathbb{L}^2(w))$ for any $T>0$, and we have a bound $\|p\|_{\mathbb{L}^\infty((0,T),\mathbb{L}^2(w))} \le C$, where C is some constant only depending on P_0, ∇V and T.

We only give the proof of Lemma 5.2.3 in the non compact case, the one in the compact case being similar.

Proof. The mild formulation (5.15) will allow us to prove that $u\in\mathbb{L}^\infty((0,T),\mathbb{L}^2(w))$. Since p_0 lies in both $\mathbb{L}^1(w)$ and $\mathbb{L}^2(w)$, it lies in $\mathbb{L}^q(w)$, for any $1\le q\le 2$. We first prove that we have a uniform in time estimate in $\mathbb{L}^q(w)$, $1\le q\le 2$, for p_t.

From equation (5.15), it follows

$$\|p_t\|_{\mathbb{L}^q(w)} \le \|p_0\|_{\mathbb{L}^q(w)} + \int_0^t \|\nabla G_{t-s} \star (\nabla V p_s)\|_{\mathbb{L}^q(w)} + \left\| \partial_1 G_{t-s} \star \left(\frac{p_s^{\partial_1 V}}{p_s^1} p_s \right) \right\|_{\mathbb{L}^q(w)} \mathrm{d}s. \tag{5.17}$$

One has, from Jensen's inequality,

$$\begin{aligned}\|\nabla G_{t-s}\star(\nabla V p_s)\|^q_{\mathbb{L}^q(w)} \leq & K\int_{\mathbb{T}\times\mathbb{R}^{d-1}} (|\nabla G_{t-s}|\star p_s)^q\, w\\ \leq & K\int_{\mathbb{T}\times\mathbb{R}^{d-1}} (|\nabla G_{t-s}|^q\star p_s)\, w\\ = & K\int_{\mathbb{T}\times\mathbb{R}^{d-1}}\int_{\mathbb{T}\times\mathbb{R}^{d-1}} |\nabla G_{t-s}(y)|^q p_s(x-y)w(x)\mathrm{d}x\mathrm{d}y.\end{aligned}$$

Now, notice that $w(x)\leq K(1+|y^{2\dots d}|^{2\lambda})w(x-y)\stackrel{\text{def}}{=}\pi(y)w(x-y)$, so that

$$\begin{aligned}\|\nabla G_{t-s}\star(\nabla V p_s)\|^q_{\mathbb{L}^q(w)} \leq & K\int_{\mathbb{T}\times\mathbb{R}^{d-1}}\int_{\mathbb{T}\times\mathbb{R}^{d-1}} |\nabla G_{t-s}(y)|^q \pi(y) p_s(x-y)w(x-y)\mathrm{d}x\mathrm{d}y\\ = & K\|p_s\|_{\mathbb{L}^1(w)}\int_{\mathbb{T}\times\mathbb{R}^{d-1}} |\nabla G_{t-s}(y)|^q\pi(y)\mathrm{d}y.\end{aligned}$$

In view of Lemma 5.1.1, $\|p_s\|_{\mathbb{L}^1(w)}$ is bounded. Moreover, one has for $0\leq s\leq t\leq T$,

$$\begin{aligned}|\nabla G_{t-s}(y)|^q\pi(y) &= \left|(4\pi(t-s))^{-d/2}\sum_{k\in\mathbb{Z}} -\frac{y-ke_1}{2(t-s)}e^{-\frac{|y-ke_1|^2}{4(t-s)}}\right|^q (1+|y^{2\dots d}|^{2\lambda})\\ &\leq \frac{K}{(t-s)^{q(d+1)/2}}\left|\left(1+\frac{|y^{2\dots d}|^{2\lambda/q}}{(t-s)^{\lambda/q}}\right)\sum_{k\in\mathbb{Z}}\frac{|y-ke_1|}{\sqrt{t-s}}e^{-\frac{|y-ke_1|^2}{4(t-s)}}\right|^q.\end{aligned}$$

Then, since a function f with polynomial growth satisfies $f(x)e^{-x^2}\leq Ke^{-x^2/2}$ for some constant K, using Hölder's inequality, we deduce,

$$\begin{aligned}|\nabla G_{t-s}(y)|^q\pi(y) &\leq \frac{K}{(t-s)^{q(d+1)/2}}\left|\sum_{k\in\mathbb{Z}} e^{-\frac{|y-ke_1|^2}{8(t-s)}}\right|^q\\ &\leq \frac{K}{(t-s)^{q(d+1)/2}}\left|\sum_{k\in\mathbb{Z}} e^{-\frac{q|y-ke_1|^2}{16(t-s)}}\right|\left|\sum_{k\in\mathbb{Z}} e^{-\frac{q'|y-ke_1|^2}{16T}}\right|^{\frac{q}{q'}}\\ &\leq \frac{K}{(t-s)^{q(d+1)/2}}\sum_{k\in\mathbb{Z}} e^{\frac{-q|y-ke_1|^2}{16(t-s)}},\end{aligned}$$

where q' satisfies $\frac{1}{q}+\frac{1}{q'}=1$. Consequently, one has, for $0\leq s\leq t\leq T$,

$$\left(\int_{\mathbb{T}\times\mathbb{R}^{d-1}}|\nabla G_{t-s}(y)|^q\pi(y)\mathrm{d}y\right)^{1/q}\leq\frac{K}{(t-s)^{(d+1)/2-d/2q}}. \tag{5.18}$$

The last term in (5.17) can be bounded in the same way, so we deduce that $\int_0^t \|\nabla G_{t-s}\star(\nabla V p_s)\|_{\mathbb{L}^q(w)}+\left\|\partial_1 G_{t-s}\star\left(\frac{p_s^{\partial_1 V}}{p_s^1}p_s\right)\right\|_{\mathbb{L}^q(w)}\mathrm{d}s$ is finite as soon as

$$1\leq q<\frac{d}{d-1}. \tag{5.19}$$

In view of (5.17), p lies in $\mathbb{L}^\infty((0,T),\mathbb{L}^q(\mathbb{T}\times\mathbb{R}^{d-1}))$ for all T and all q satisfying (5.19), and we have a bound on its norm depending only on P_0, ∇V and T. We now bootstrap this estimate to reach a uniform-in-time $\mathbb{L}^2(w)$ bound for p.

Let n_0 be an integer large enough so that

$$\frac{n_0+1}{n_0+1/2} < \frac{d}{d-1},$$

and define for $n = 0, \dots n_0$, $q = \frac{n_0+1}{n_0+1/2}$ and $q_n = \left(\frac{1}{q} + n\left(\frac{1}{q} - 1\right)\right)^{-1}$. Notice that $(q_n)_{n=0\dots n_0}$ satisfies $q_0 = q$, $q_{n_0} = 2$ and

$$1 + \frac{1}{q_{n+1}} = \frac{1}{q_n} + \frac{1}{q},$$

so that, according to Young's Inequality, convolution continuously maps $\mathbb{L}^{q_n} \times \mathbb{L}^q$ to $\mathbb{L}^{q_{n+1}}$. Consequently, we have for $n < n_0$

$$\begin{aligned}\|\nabla G_{t-s} \star (\nabla V p_s)\|_{\mathbb{L}^{q_{n+1}}(w)} \leq & K\left(\int_{\mathbb{T}\times\mathbb{R}^{d-1}} (|\nabla G_{t-s}| \star p_s)^{q_{n+1}}(x) w(x) \mathrm{d}x\right)^{1/q_{n+1}} \\ = & K\left(\int_{\mathbb{T}\times\mathbb{R}^{d-1}} \left(\int_{\mathbb{T}\times\mathbb{R}^{d-1}} |\nabla G_{t-s}|(y) p_s(x-y)\mathrm{d}y\right)^{q_{n+1}} w(x)\mathrm{d}x\right)^{1/q_{n+1}}.\end{aligned}$$

We have $w(x) \leq w(x-y)\pi(y) \leq w(x-y)^{q_{n+1}/q_n}\pi(y)$, since $q_n \leq q_{n+1}$, yielding, by Young's inequality and the polynomial growth of π,

$$\begin{aligned}\|\nabla G_{t-s} \star (\nabla V p_s)\|_{\mathbb{L}^{q_{n+1}}(w)} \leq & K\left(\int_{\mathbb{T}\times\mathbb{R}^{d-1}} \left(\int_{\mathbb{T}\times\mathbb{R}^{d-1}} |\nabla G_{t-s}|(y)\pi(y)^{1/q_{n+1}}\right.\right. \\ & \left.\left. \times p_s(x-y) w(x-y)^{1/q_n}\mathrm{d}y\right)^{q_{n+1}} \mathrm{d}x\right)^{1/q_{n+1}} \\ = & K\|(|\nabla G_{t-s}|\pi^{1/q_{n+1}}) \star (p_s w^{1/q_n})\|_{\mathbb{L}^{q_{n+1}}(\mathbb{T}\times\mathbb{R}^{d-1})} \\ \leq & K\||\nabla G_{t-s}|\pi^{1/q_{n+1}}\|_{\mathbb{L}^q(\mathbb{T}\times\mathbb{R}^{d-1})}\|p_s\|_{\mathbb{L}^{q_n}(w)} \\ \leq & \frac{K}{(t-s)^{(d+1)/2-d/(2q)}}\|p_s\|_{\mathbb{L}^{q_n}(w)},\end{aligned}$$

the last inequality being proved in the same way as (5.18) is. As a result, for $n < n_0$,

$$\|p_t\|_{\mathbb{L}^{q_{n+1}}(w)} \leq \|p_0\|_{\mathbb{L}^{q_{n+1}}(w)} + K\int_0^t \frac{\|p_s\|_{\mathbb{L}^{q_n}(w)}}{(t-s)^{(d+1)/2-d/(2q)}}\mathrm{d}s.$$

By induction on n, since $\frac{1}{(t-s)^{(d+1)/2-d/(2q)}}$ is integrable on $[0,t]$, this estimate shows that p lies in $\mathbb{L}^\infty((0,T),\mathbb{L}^2(w))$, for all positive T. Since we control $\sup_{t\in[0,T]}\|p_t\|_{\mathbb{L}^{q_0}(w)}$ by a constant depending only on P_0, ∇V and T, we also have such a control on $\sup_{t\in[0,T]}\|p_t\|_{\mathbb{L}^2(w)}$.

Now, we prove that p is a solution to Equation (5.12) in the sense of Definition 5.2.1. First, we show that it satisfies the regularity condition.

Lemma 5.2.4 *In the compact case, under Assumptions i and ii, one has*

$$p \in \mathbb{L}^\infty\left((0,T),\mathbb{L}^2(\mathbb{T}^d)\right) \bigcap \mathbb{L}^2\left((0,T),\mathbb{H}^1(\mathbb{T}^d)\right). \tag{5.20}$$

Moreover $\|p\|_{\mathbb{L}^\infty((0,T),\mathbb{L}^2(\mathbb{T}^d))} + \|p\|_{\mathbb{L}^2((0,T),\mathbb{H}^1(\mathbb{T}^d))} \leq K$, *where* K *only depends on* ∇V, P_0 *and* T.

In the non compact case, with the additional Assumption iii, one has

$$p \in \mathbb{L}^\infty\left((0,T),\mathbb{L}^2(w)\right) \bigcap \mathbb{L}^2\left((0,T),\mathbb{H}^1(w)\right). \tag{5.21}$$

Moreover $\|p\|_{\mathbb{L}^\infty((0,T),\mathbb{L}^2(w))} + \|p\|_{\mathbb{L}^2((0,T),\mathbb{H}^1(w))} \leq K$, *where* K *only depends on* ∇V, P_0 *and* T.

Proof. According to Assumption ii, p_0 lies in $\mathbb{L}^2(\mathcal{D})$. Consequently, from Lemma 5.2.3, we know that P_t has a density p_t such that $p \in \mathbb{L}^\infty((0,T),\mathbb{L}^2(\mathcal{D}))$. We now prove that p lies in $\mathbb{L}^2((0,T),\mathbb{H}^1(\mathcal{D}))$. We know that p lies in $\mathbb{L}^\infty((0,T),\mathbb{L}^2(\mathcal{D})) \subset \mathbb{L}^2((0,T),\mathbb{L}^2(\mathcal{D}))$, and that $\frac{p^{\partial_1 V}}{p^1}$ is in $\mathbb{L}^\infty([0,T]\times\mathcal{D})$, so that the function f defined by

$$f = \operatorname{div}(p\nabla V) - \partial_1\left(\frac{p^{\partial_1 V}}{p^1}p\right)$$

lies in $\mathbb{L}^2((0,T),\mathbb{H}^{-1}(\mathcal{D}))$. Consequently, it can be shown, for example using a Galerkin approximation (see [26, Chapter XVIII]) that the problem

$$\begin{cases} \partial_t v - \Delta v = f, \\ v_0 = p_0, \end{cases} \tag{5.22}$$

admits a unique weak solution v in the space $\mathbb{L}^\infty((0,T),\mathbb{L}^2(\mathcal{D}))\bigcap\mathbb{L}^2((0,T),\mathbb{H}^1(\mathcal{D}))$. Here, "weak solution" means that for any ψ in $\mathbb{H}^1(\mathcal{D})$,

$$\partial_t \int_{\mathcal{D}} \psi v_t + \int_{\mathcal{D}} \nabla\psi\nabla v_t = \int_{\mathcal{D}} \psi f \tag{5.23}$$

holds. Thanks to an *a priori* estimate, we can find a bound K depending only on ∇V, P_0 and T, such that this weak solution lies in the ball of radius C in the spaces $\mathbb{L}^\infty((0,T),\mathbb{L}^2(\mathcal{D}))$ and $\mathbb{L}^2((0,T),\mathbb{H}^1(\mathcal{D}))$. For the non compact case, notice that under Assumption iii, f satisfies for any $\psi \in \mathbb{H}^1(w)$,

$$\begin{aligned}\left|\int_{\mathbb{T}\times\mathbb{R}^{d-1}} f\psi w\right| &= \left|\int_{\mathbb{T}\times\mathbb{R}^{d-1}} p\nabla V\cdot\nabla(\psi w) - \frac{p^{\partial_1 V}}{p^1}p\,\partial_1(\psi w)\right| \\ &\le K\int_{\mathbb{T}\times\mathbb{R}^{d-1}} |p\nabla\psi|w + K\int_{\mathbb{T}\times\mathbb{R}^{d-1}} |p\psi|w \\ &\le K\|\psi\|_{\mathbb{H}^1(w)},\end{aligned}$$

the last bound being deduced from Lemma 5.2.3. From the following *a priori* estimate,

$$\begin{aligned}\frac{1}{2}\partial_t\|v_t\|^2_{\mathbb{L}^2(w)} &= -\int_{\mathbb{T}\times\mathbb{R}^{d-1}} \nabla v_t\nabla(wv_t) + \int_{\mathbb{T}\times\mathbb{R}^{d-1}} f v_t w \\ &\le -\int_{\mathbb{T}\times\mathbb{R}^{d-1}} |\nabla v_t|^2 w + K\int_{\mathbb{T}\times\mathbb{R}^{d-1}} |v_t\nabla v_t|w + K\|v_t\|_{\mathbb{H}^1(w)} \\ &\le -\frac{1}{2}\|\nabla v_t\|^2_{\mathbb{L}^2(w)} + K\|v_t\|^2_{\mathbb{L}^2(w)} + K,\end{aligned}$$

standard arguments show that v also lies in $\mathbb{L}^\infty((0,T),\mathbb{L}^2(w))\bigcap\mathbb{L}^2((0,T),\mathbb{H}^1(w))$, if $p_0 \in \mathbb{L}^2(w)$.

We are now going to show that v is actually equal to the function p. For a fixed t in $[0,T]$, consider $\psi_s = G_{t-s}\star\chi$, solution to the problem (5.16), where χ is some test function, and compute $\partial_s\int_{\mathcal{D}}\psi_s v_s$. From [67, page 261, Lemma 1.2], we obtain

$$\partial_s\int_{\mathcal{D}}\psi_s v_s = \int_{\mathcal{D}}\psi_s f,$$

in the sense of distributions. Using the expression of ψ_s, this equation rewrites

$$\partial_s\int_{\mathcal{D}}(G_{t-s}\star\chi)v_s = \int_{\mathcal{D}}(G_{t-s}\star\chi)f,$$

which is equivalent to

$$\partial_s \int_{\mathcal{D}} \chi(G_{t-s} \star v_s) = \int_{\mathcal{D}} \chi(G_{t-s} \star f). \tag{5.24}$$

Since $v \in \mathbb{L}^2((0,T), \mathbb{H}^1(\mathcal{D}))$, and $\partial_s v \in \mathbb{L}^2((0,T), \mathbb{H}^{-1}(\mathcal{D}))$, then v lies in $\mathcal{C}((0,T), \mathbb{L}^2(\mathcal{D}))$ (see [26, Chapter XVIII, §1, Theorem 1], so that the left hand side in (5.24) is the derivative with respect ot s of a function which is continuous in s. Moreover, one has

$$G_{t-s} \star f = \nabla G_{t-s} \star (p \nabla V) - \partial_1 G_{t-s} \star \left(\frac{p^{\partial_1 V}}{p^1} p \right) \in \mathbb{L}^1((0,t), \mathbb{L}^2(\mathcal{D})),$$

so that the right hand side in (5.24) is integrable in time. Consequently, integrating on $[0,t]$, one finds

$$\int_{\mathcal{D}} \chi v_t = \int_{\mathcal{D}} \chi(G_t \star p_0) + \int_{\mathcal{D}} \int_0^t \chi(\nabla G_{t-s} \star (\nabla V p_s)) \mathrm{d}s - \int_{\mathcal{D}} \int_0^t \chi \left(\partial_1 G_{t-s} \star \left(\frac{p_s^{\partial_1 V}}{p_s^1} p_s \right) \right) \mathrm{d}s.$$

Identifying in the sense of distribution, one has

$$v_t = G_t \star p_0 + \int_0^t \nabla G_{t-s} \star (\nabla V p_s) \mathrm{d}s - \int_0^t \partial_1 G_{t-s} \star \left(\frac{p_s^{\partial_1 V}}{p_s^1} p_s \right) \mathrm{d}s. \tag{5.25}$$

The right hand side in (5.25) is exactly the right hand side in (5.15), and (5.25) holds for all $t > 0$, so that $v = p$, and the regularity we wanted on p actually holds.

We finish this section by proving :

Lemma 5.2.5 *The function p satisfies Equation* (5.12) *in the sense of Definition 5.2.1.*

Proof. According to Lemma 5.2.4, in the compact case (resp. in the non compact case), for any $t > 0$, p lies in $\mathbb{L}^\infty((0,T), \mathbb{L}^2(\mathbb{T}^d) \bigcap \mathbb{L}^2((0,T), \mathbb{H}^1(\mathbb{T}^d))$ (resp. in $\mathbb{L}^\infty((0,T), \mathbb{L}^2(w) \bigcap \mathbb{L}^2((0,T), \mathbb{H}^1(w))$). Moreover, thanks to Itō's Formula, p satisfies Equation (5.11) for any smooth test function ψ. But, according to the regularity of p_t, and by the density of smooth functions in $\mathbb{H}^1(\mathbb{T}^d)$ (resp. in $\mathbb{H}^1(w)$), Equation (5.13) holds for any ψ in $\mathbb{H}^1(\mathbb{T}^d)$ (resp. (5.14) holds for any ψ in $\mathbb{H}^1(w)$). This means that p_t is a solution to (5.12) in the sense of Definition 5.2.1.

5.2.2 Uniqueness results

In this section we prove that uniqueness holds for solutions of Equation (5.12) in the sense of Definition 5.2.1, yielding uniqueness for solutions of the nonlinear equation (5.6).

Uniqueness for the Fokker-Planck Equation

Theorem 5.2.6 *In the compact case, under Assumptions i and ii or in the non compact case under Assumptions i, ii and iii, uniqueness holds for the solutions to the Fokker-Planck equation* (5.12) *in the sense of Definition 5.2.1.*

Proof. We only give the proof in the non compact case, which can be adapted straightforwardly for the compact case by performing the same computations in the space $\mathbb{L}^2(\mathbb{T}^d)$. Let u and v be two solutions of (5.12) in the sense of Definition 5.2.1 with same initial condition $u_0 = v_0$. We use Grönwall's Lemma to prove that $\|u_t - v_t\|_{\mathbb{L}^2(w)} = 0$ for all $t > 0$. Adapting the proof of [67, page 261, Lemma 1.2], one has $\frac{1}{2}\partial_t \|u_t - v_t\|^2_{\mathbb{L}^2(w)} = \int_{\mathbb{T}\times\mathbb{R}^{d-1}} (u_t - v_t)\partial_t(u_t - v_t)w$. Consequently, since u and v satisfy Definition 5.2.1, and using (5.5) and Assumption i, it holds that

$$\frac{1}{2}\partial_t \|u_t - v_t\|^2_{\mathbb{L}^2(w)} \leq K\|u_t - v_t\|^2_{\mathbb{L}^2(w)} + K\|u_t - v_t\|_{\mathbb{L}^2(w)} \|\nabla u_t - \nabla v_t\|_{\mathbb{L}^2(w)} - \|\nabla u_t - \nabla v_t\|^2_{\mathbb{L}^2(w)}$$
$$+ \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \partial_1(u_t - v_t) \left(u_t \frac{u_t^{\partial_1 V}}{u_t^1} - v_t \frac{v_t^{\partial_1 V}}{v_t^1} \right) w.$$

We want to estimate the last term. Notice that, thanks to Lemma 5.1.2, $u^1 = v^1$, so that

$$\int_{\mathbb{T}\times\mathbb{R}^{d-1}} \partial_1(u_t - v_t)\left(u_t\frac{u_t^{\partial_1 V}}{u_t^1} - v_t\frac{v_t^{\partial_1 V}}{v_t^1}\right)w = \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \partial_1(u_t - v_t)u_t\frac{u_t^{\partial_1 V} - v_t^{\partial_1 V}}{u_t^1}w + \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \partial_1(u_t - v_t)(u_t - v_t)\frac{v_t^{\partial_1 V}}{u_t^1}w.$$

Since $\partial_1 V$ is bounded, the second term in the right-hand side is smaller than

$$K\|u_t - v_t\|_{\mathbb{L}^2(w)}\|\nabla u_t - \nabla v_t\|_{\mathbb{L}^2(w)},$$

and the first term is smaller than

$$\|\nabla u_t - \nabla v_t\|_{\mathbb{L}^2(w)}\left(\int_{\mathbb{T}\times\mathbb{R}^{d-1}}\left(\frac{u_t}{u_t^1}\right)^2\left(u_t^{\partial_1 V} - v_t^{\partial_1 V}\right)^2 w\right)^{1/2}.$$

Then,

$$\left(\int_{\mathbb{T}\times\mathbb{R}^{d-1}}\left(\frac{u_t}{u_t^1}\right)^2\left(u_t^{\partial_1 V} - v_t^{\partial_1 V}\right)^2 w\right)^{1/2} = \left(\int_{\mathbb{T}}\left(\frac{u_t^{\partial_1 V} - v_t^{\partial_1 V}}{u_t^1}\right)^2\left(\int_{\mathbb{R}^{d-1}}(u_t)^2 w\right)\right)^{1/2} \leq \left\|\frac{1}{u_t^1}(u_t^{\partial_1 V} - v_t^{\partial_1 V})\right\|_{\mathbb{L}^\infty(\mathbb{T})}\|u_t\|_{\mathbb{L}^2(w)}.$$

The function $t \mapsto \|u_t\|_{\mathbb{L}^2(w)}$ is bounded on $[0,T]$, and, thanks to Lemma 5.1.2, Assumption ii and the maximum principle, u^1 is bounded from below by some positive constant, so that

$$\left(\int_{\mathbb{T}\times\mathbb{R}^{d-1}}\left(\frac{u_t}{u_t^1}\right)^2\left(u_t^{\partial_1 V} - v_t^{\partial_1 V}\right)^2 w\right)^{1/2} \leq K\|u_t^{\partial_1 V} - v_t^{\partial_1 V}\|_{\mathbb{L}^\infty(\mathbb{T})}.$$

To conclude, notice that, for any positive γ, $\mathbb{H}^{1/2+\gamma}(\mathbb{T})$ continuously imbeds in $\mathcal{C}(\mathbb{T})$ (see [1, page 217]). Consequently, interpolating $\mathbb{H}^1(\mathbb{T})$ and $\mathbb{L}^2(\mathbb{T})$ (see [51, Page 49], we obtain for a function f in $\mathbb{H}^1(\mathbb{T})$ and $\gamma \in (0, \frac{1}{2})$,

$$\|f\|_{\mathbb{L}^\infty(\mathbb{T})} \leq K\|f\|_{\mathbb{H}^{1/2+\gamma}(\mathbb{T})} \leq K\|f\|_{\mathbb{L}^2(\mathbb{T})}^{1/2-\gamma}\|f\|_{\mathbb{H}^1(\mathbb{T})}^{1/2+\gamma}. \tag{5.26}$$

All the previous inequalities give us

$$\frac{1}{2}\partial_t\|u_t - v_t\|_{\mathbb{L}^2(w)}^2 + \|\nabla u_t - \nabla v_t\|_{\mathbb{L}^2(w)}^2 \leq K\|u_t - v_t\|_{\mathbb{L}^2(w)}\|\nabla u_t - \nabla v_t\|_{\mathbb{L}^2(w)} + K\|u_t - v_t\|_{\mathbb{L}^2(w)}^2 + K\|u_t^{\partial_1 V} - v_t^{\partial_1 V}\|_{\mathbb{L}^2(\mathbb{T})}^{1/2-\gamma}\|u_t^{\partial_1 V} - v_t^{\partial_1 V}\|_{\mathbb{H}^1(\mathbb{T})}^{1/2+\gamma}\|\nabla u_t - \nabla v_t\|_{\mathbb{L}^2(w)}.$$

We finally obtain, from Lemma 5.0.1 and Young's inequality $ab \leq \varepsilon a^p + q^{-1}(p\varepsilon)^{-q/p}b^q$, holding true for any positive a, b, ε, p and q such that $\frac{1}{p} + \frac{1}{q} = 1$,

$$\partial_t\|u_t - v_t\|_{\mathbb{L}^2(w)}^2 + \|\nabla u_t - \nabla v_t\|_{\mathbb{L}^2(w)}^2 \leq K\|u_t - v_t\|_{\mathbb{L}^2(w)}^2,$$

yielding uniqueness through Grönwall's lemma.

Remark 5.2.7 *A more natural uniqueness proof can be performed, using an entropy estimate. In particular, this proof does not require the introduction of the weighted spaces. Unfortunately, it*

does not apply to the solutions in the sense of Definition 5.2.1. Uniqueness actually holds in the subspace of functions such that the following computations make sense.

Let u and v be two solutions of (5.12) with same initial condition $u_0 = v_0$. Notice that from Lemma 5.1.2, the functions u^1 and v^1 are equal. Define the relative entropy of u with respect to v :

$$E(t) = \int_{\mathbb{T}\times\mathbb{R}^{d-1}} u \log \frac{u}{v}.$$

If all quantities involved are finite, it holds that

$$\begin{aligned}
E'(t) &= \partial_t \left(\int_{\mathbb{T}\times\mathbb{R}^{d-1}} u \right) + \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \partial_t u \ \log \frac{u}{v} - \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \partial_t v \ \frac{u}{v} \\
&= 0 - \int_{\mathbb{T}\times\mathbb{R}^{d-1}} u \nabla V \ \nabla \log \frac{u}{v} - \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \nabla u \ \nabla \log \frac{u}{v} + \int_{\mathbb{T}\times\mathbb{R}^{d-1}} u \frac{u^{\partial_1 V}}{u^1} \partial_1 \log \frac{u}{v} \\
&\quad + \int_{\mathbb{T}\times\mathbb{R}^{d-1}} v \nabla V \ \nabla \frac{u}{v} + \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \nabla v \ \nabla \frac{u}{v} - \int_{\mathbb{T}\times\mathbb{R}^{d-1}} v \frac{v^{\partial_1 V}}{u^1} \partial_1 \frac{u}{v} \\
&= - \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \frac{v^2}{u} \left| \nabla \frac{u}{v} \right|^2 + \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \left(u^{\partial_1 V} - v^{\partial_1 V} \right) \ \frac{v}{u^1} \partial_1 \frac{u}{v}.
\end{aligned}$$

But, using Csiszár-Kullback inequality, it holds that

$$\begin{aligned}
\int_{\mathbb{T}\times\mathbb{R}^{d-1}} \left(u^{\partial_1 V} - v^{\partial_1 V} \right) \ \frac{v}{u^1} \partial_1 \frac{u}{v} &\leq K \int_{\mathbb{T}\times\mathbb{R}^{d-1}} v \left| \partial_1 \frac{u}{v} \right| \left\| \frac{u}{u^1} - \frac{v}{u^1} \right\|_{TV(\mathbb{R}^{d-1})} \\
&\leq K \int_{\mathbb{T}\times\mathbb{R}^{d-1}} v \left| \partial_1 \frac{u}{v} \right| \left(\int_{\mathbb{R}^{d-1}} \left(\frac{u}{u^1} \log \frac{u}{v} \right) \right)^{1/2}.
\end{aligned}$$

In conclusion, we find

$$E'(t) \leq - \int_{\mathbb{T}\times\mathbb{R}^{d-1}} \frac{v^2}{u} \left| \nabla \frac{u}{v} \right|^2 + K \left(\int_{\mathbb{T}\times\mathbb{R}^{d-1}} \frac{v^2}{u} \left| \partial_1 \frac{u}{v} \right|^2 \right)^{1/2} (E(t))^{1/2}.$$

We can conclude the proof using Young's inequality and then Grönwall's Lemma.

Uniqueness for the nonlinear process

Theorem 5.2.8 *Pathwise uniqueness and uniqueness in law hold for Equation (5.6) in the compact case under Assumptions i and ii, and in the non compact case under Assumptions i, ii and iii.*

Proof. As stated in Lemma 5.2.5, if X solves (5.6), then $\{X_t\}$ admits a density p_t such that p satisfies (5.12) in the sense of Definition 5.2.1. Thus, in regard of Theorem 5.2.6, p_t is uniquely defined. Consequently, Equation (5.6) rewrites

$$\mathrm{d}X_t = \left(-\nabla V(X_t) + \frac{p_t^{\partial_1 V}(X_t^1)}{p_t^1(X_t^1)} e_1 \right) \mathrm{d}t + \sqrt{2} \mathrm{d}W_t, \tag{5.27}$$

where p_t is the unique solution to Equation (5.12) in the sense of Definition 5.2.1. Notice that the drift

$$b_t(x) = -\nabla V(x) + \frac{p_t^{\partial_1 V}(x^1)}{p_t^1(x^1)} e_1$$

in Equation (5.27) is bounded, so that pathwise uniqueness holds (see [45]), as well as uniqueness in law, from the Girsanov Theorem.

5.3 A regularized approximate dynamics

To estimate the difference between the nonlinear process defined by Equation (5.6) and its particle approximation (5.8), we introduce an intermediate process, called the regularized nonlinear process, which is the natural expected limit as N goes to infinity of the particle approximation (5.8). The nonlinear term in this equation is more regular than the one in (5.6), so that we can show existence and uniqueness for this process.

The aim of this section will be, in a first time, to prove existence and uniqueness for the regularized nonlinear process, see Theorem 5.3.1, and in a second time to show that the regularized nonlinear process converges to the nonlinear process solution to (5.6) as ε and α go to zero, and to estimate the rate of this convergence, see Theorem 5.3.11 below. This will yield an existence result for the nonlinear process.

Under Assumption iv on the initial condition, for a fixed positive integer n, we expect the sequence of processes $(X^{\eta}_{n,N})_{N>0}$ defined by (5.8) to converge to a solution to

$$\begin{cases} \mathrm{d}\bar{X}^{\eta}_{t,n} &= \left(-\nabla V(\bar{X}^{\eta}_{t,n}) + \dfrac{\varphi_\eta * P^{\eta,\partial_1 V}_t}{\varphi_\eta * P^{\eta,1}_t}(\bar{X}^{\eta,1}_{t,n})e_1\right)\mathrm{d}t + \sqrt{2}\mathrm{d}W^n_t, \\ P^{\eta}_t &= \text{distribution of } \{\bar{X}^{\eta}_{t,n}\} \end{cases} \tag{5.28}$$

with initial condition $(X_{0,n})$.

5.3.1 Existence and uniqueness for the regularized problem

In this section, we show that pathwise uniqueness, uniqueness in distribution and strong existence hold for the regularized dynamics.

We first show existence and uniqueness of a solution to (5.28), using a fixed point method.

Theorem 5.3.1 *Consider both the compact and the non compact cases. Under Assumptions i and iv, strong existence and uniqueness hold for Equation* (5.28).

Here we follow [64] : we show that a measure on the space of continuous paths from $[0,T]$ to $\mathbb{R}^d$ is the law of a solution to (5.28) if and only if it is a fixed point of some function Φ_T. Then we show existence and uniqueness of this fixed point by a contraction argument. This cannot be done directly for Equation (5.6), since its nonlinear term is too ill-behaved, so that we do not have contraction in that case.

For a probability measure μ on the set $\mathcal{C}_T = \mathcal{C}([0,T],\mathbb{R}^d)$ we denote by $\Phi_T(\mu)$ the distribution on $\mathcal{C}_T$ of the process X defined by

$$\mathrm{d}X_t = \left(-\nabla V(X_t) + \frac{\displaystyle\int_{\mathcal{C}_T} \varphi_\eta(X^1_t - x^1_t)\partial_1 V(x_t)\mathrm{d}\mu(x)}{\displaystyle\int_{\mathcal{C}_T} \varphi_\eta(X^1_t - x^1_t)\mathrm{d}\mu(x)} e_1\right)\mathrm{d}t + \sqrt{2}\mathrm{d}W_t \tag{5.29}$$

whose initial condition X_0 has law P_0 and is independent of W. The distribution $\Phi_T(\mu)$ is well defined since Equation (5.29), having global Lipschitz coefficients, has a unique strong solution.

Notice that, since

$$\frac{\varphi_\eta * \mu^{\partial_1 V}_t}{\varphi_\eta * \mu^1_t} = \frac{\int_{\mathcal{C}_T} \varphi_\eta(\cdot - x^1_t)\partial_1 V(x_t)\mathrm{d}\mu(x)}{\int_{\mathcal{C}_T} \varphi_\eta(\cdot - x^1_t)\mathrm{d}\mu(x)},$$

μ is the distribution of a solution to (5.28) up to time T if, and only if $\Phi_T(\mu) = \mu$. We will show that such a μ exists and is unique using Picard's Theorem.

The Wasserstein metric $D_T(\mu_1, \mu_2)$ between two probability distributions μ_1 and μ_2 on $\mathcal{C}_T$ is defined by

$$D_T(\mu_1, \mu_2) = \inf_{\pi \in \Pi} \int_{\mathcal{C}_T \times \mathcal{C}_T} 1 \wedge \|x - y\|_{\mathcal{C}_T} \mathrm{d}\pi(x, y),$$

where $\Pi = \{\pi \in \mathcal{P}(\mathcal{C}_T \times \mathcal{C}_T),\ \pi \textit{ having } \mu_1 \textit{ and } \mu_2 \textit{ as marginal distributions}\}$ is the set of all coupling of μ_1 and μ_2, and $\|.\|_{\mathcal{C}_T}$ is the uniform norm on $\mathcal{C}_T$:

$$\|f - g\|_{\mathcal{C}_T} = \sup_{t \in [0,T]} |f(t) - g(t)|.$$

More generally, for $t \in [0, T]$, we set

$$D_t(\mu_1, \mu_2) = \inf_{\pi \in \Pi} \int_{\mathcal{C}_T \times \mathcal{C}_T} 1 \wedge \sup_{s \in [0,t]} |x_s - y_s| \mathrm{d}\pi(x, y).$$

Endowed with the Wasserstein metric, the space $\mathcal{P}(\mathcal{C}_T)$ of probability measures on $\mathcal{C}_T$ is complete. In order to apply a fixed point argument, we will need the following contraction lemma.

Lemma 5.3.2 *Consider both the compact and non compact case. Let T be a positive time. Under Assumption i, there is a positive constant K, not depending on t, satisfying*

$$D_t(\Phi_T(\mu_1), \Phi_T(\mu_2)) \leq K \int_0^t D_s(\mu_1, \mu_2) \mathrm{d}s,$$

for all t in $[0, T]$ and for all probability measures μ_1 and μ_2 in $\mathcal{P}(\mathcal{C}_T)$.

Proof. Let μ_1 and μ_2 be two probability measures on $\mathcal{C}_T$. For $i = 1, 2$, define $X_{t,i}$ by

$$\mathrm{d}X_{t,i} = \left(-\nabla V(X_{t,i}) + \frac{\displaystyle\int_{\mathcal{C}_T} \varphi_\eta(X^1_{t,i} - x^1_t) \partial_1 V(x_t) \mathrm{d}\mu_i(x)}{\displaystyle\int_{\mathcal{C}_T} \varphi_\eta(X^1_{t,i} - x^1_t) \mathrm{d}\mu_i(x)} e_1 \right) \mathrm{d}t + \sqrt{2} \mathrm{d}W_t$$

with given initial condition $X_{0,i} = X_0$, for $i = 1, 2$.

Notice that

$$\frac{\int_{\mathcal{C}_T} \varphi_\eta(\cdot - x^1_t) \partial_1 V(x_t) \mathrm{d}\mu_i(x)}{\int_{\mathcal{C}_T} \varphi_\eta(\cdot - x^1_t) \mathrm{d}\mu_i(x)} = \frac{\varphi_\eta * \mu^{\partial_1 V}_{i,t}}{\varphi_\eta * \mu^1_{i,t}}, \tag{5.30}$$

and that from (5.9) and Assumption i, the numerator and the denominator of (5.30) are respectively bounded from above and from below by positive constants depending only on η and V. Then, for any x, y and $0 \leq s \leq T$,

$$\left| \frac{\varphi_\eta * \mu^{\partial_1 V}_{1,s}}{\varphi_\eta * \mu^1_{1,s}}(x) - \frac{\varphi_\eta * \mu^{\partial_1 V}_{2,s}}{\varphi_\eta * \mu^1_{2,s}}(y) \right| \leq K \left(|x - y| \wedge 1 + D_s(\mu_1, \mu_2) \right).$$

Consequently,

$$\mathbb{E}\left[1 \wedge \|X_1 - X_2\|_{\mathcal{C}_t}\right] \leq K \left(\int_0^t \mathbb{E}\left[1 \wedge \|X_1 - X_2\|_{\mathcal{C}_s}\right] \mathrm{d}s + \int_0^t D_s(\mu_1, \mu_2) \mathrm{d}s \right),$$

for all $t \leq T$. Using Grönwall's Lemma, we then find, for any $t \leq T$,

$$\mathbb{E}\left[1 \wedge \|X_1 - X_2\|_{\mathcal{C}_t}\right] \leq K \int_0^t D_s(\mu_1, \mu_2)\mathrm{d}s.$$

But

$$D_t(\Phi_t(\mu_1), \Phi_t(\mu_2)) \leq \mathbb{E}\left[1 \wedge \|X_1 - X_2\|_{\mathcal{C}_t}\right]$$

since X_1 and X_2 respectively have $\Phi_t(\mu_1)$ and $\Phi_t(\mu_2)$ as distributions, finishing the proof.

Proof (Proof of Theorem 5.3.1). Iterating Lemma 5.3.2, we find existence and uniqueness of a fixed point of Φ_T, given X_0, which yields uniqueness of the distribution P of the solution to (5.28) on $[0, T]$.

The law P of any solution being unique, we can substitute the marginal of P at time t in Equation (5.28), and we obtain a linear stochastic differential equation with Lipschitz continuous coefficients. Pathwise uniqueness holds for that kind of equation, so that weak existence and pathwise uniqueness hold for (5.28). Consequently, from Yamada-Watanabe Theorem, it admits a unique strong solution.

5.3.2 Convergence to the nonlinear process

We are now going to let ε and α go to 0 in (5.28).

We denote by X_t^η the unique strong solution to (5.28), with initial condition X_0 and Brownian motion W^n replaced with W. The distribution of $\{X_t^\eta\}$ will be denoted P^η. We expect a possible limit X of X^η as η goes to 0 to be a solution to (5.6). To this aim, we define the following martingale problem :

Definition 5.3.3 *We say that a probability measure P on the space $\mathcal{C}_T$ of continuous paths is a solution to the martingale problem associated to* (5.6) *if its time marginals P_t admit a density p_t with respect to the Lebesgue measure, and if, under the measure P,*

- *the canonical process $x \in \mathcal{C}_T$ is such that for any twice differentiable function which is bounded as well as its first and second derivatives, the process*

$$m_t = \psi(x_t) - \psi(x_0) + \int_0^t \nabla\psi(x_s)\nabla V(x_s)\mathrm{d}s - \int_0^t \Delta\psi(x_s)\mathrm{d}s - \int_0^t \partial_1\psi(x_s)\frac{p_t^{\partial_1 V}(x_t)}{p_t^1(x_t)}\mathrm{d}s, \quad (5.31)$$

  is a martingale with respect to the filtration $\sigma(x_s, s \leq t)$.
- *$\{x_0\}$ has law P_0.*

Notice that, since the drift coefficient is bounded, the Girsanov theorem shows that it is not restrictive to assume that P_t has a density.

We deduce from Theorem 5.2.8 the following result :

Proposition 5.3.4 *In the compact case under Assumptions i and ii, or in the non compact case under Asumptions i, ii and iii, uniqueness holds for the martingale problem defined in Definition 5.3.3.*

Our aim in this section will be to prove the following results :

Theorem 5.3.5 *Let Assumptions i and ii hold.*

In the compact case, $(P^\eta)_{\eta>0}$ converges as η goes to 0 to the solution of the martingale problem.

In the non compact case, the family of probability measures $(P^\eta)_{\eta>0}$ is tight, and any converging subsequence converges to a solution of the martingale problem defined in Definition 5.3.3. Under the additional Assumption iii, $(P_\eta)_{\eta>0}$ actually converges to the unique solution.

As a corollary of Theorem 5.3.5, one has existence of solutions to (5.6) (under regularity assumptions on the initial condition).

From Proposition 5.3.4, in order to prove Theorem 5.3.5, it is enough to prove that the family $(P^\eta)_{\eta>0}$ is tight, and that any converging subsequence converges to a solution of the martingale problem.

Our first step will be to derive the Fokker-Planck equation satisfied by the distribution of $\{X_t^\eta\}$. Let ψ be a smooth bounded function on $\mathcal{D}$, with bounded derivatives. Applying Itō's formula to $\psi(X_t^\eta)$ and taking the expectation, we find that

$$\int_{\mathcal{D}} \psi \mathrm{d}P_T^\eta = \int_{\mathcal{D}} \psi p_0(x)\mathrm{d}x + \int_0^T \int_{\mathcal{D}} \left(\Delta\psi - \nabla\psi\cdot\nabla V\right) \mathrm{d}P_t^\eta \mathrm{d}t + \int_0^T \int_{\mathcal{D}} \partial_1\psi \frac{\varphi_\eta * P_t^{\eta,\partial_1 V}}{\varphi_\eta * P_t^{\eta,1}} \mathrm{d}P_t^\eta \mathrm{d}t. \tag{5.32}$$

Equation (5.32) is a weak formulation of the following partial differential equation

$$\partial_t P_t^\eta = \mathrm{div}\left(P_t^\eta \nabla V + \nabla P_t^\eta\right) - \partial_1 \left(P_t^\eta \frac{\varphi_\eta * P_t^{\eta,\partial_1 V}}{\varphi_\eta * P_t^{\eta,1}} \right). \tag{5.33}$$

We are going to show that P_t^η, or more precisely, its density, is actually a solution to equation (5.33) in the following stronger sense.

Definition 5.3.6 *A function u is said to be a solution to* (5.33) *with initial condition p_0 if, in the compact case,*

- *u belongs to $\mathbb{L}^\infty((0,T),\mathbb{L}^2(\mathbb{T}^d)) \bigcap \mathbb{L}^2((0,T),\mathbb{H}^1(\mathbb{T}^d))$;*
- *for any function $\psi \in \mathbb{H}^1(\mathbb{T}^d)$, we have :*

$$\partial_t \int_{\mathcal{D}} u_t\psi = -\int_{\mathcal{D}} u_t \nabla V\cdot\nabla\psi - \int_{\mathcal{D}} \nabla u_t\cdot\nabla\psi + \int_{\mathcal{D}} (\partial_1\psi) u_t \frac{\varphi_\eta * u_t^{\partial_1 V}}{\varphi_\eta * u_t^1} \tag{5.34}$$

  in the sense of distributions in time ;
- *$u_0 = p_0$.*

In the non compact case these conditions are replaced by

- *u belongs to $\mathbb{L}^\infty((0,T),\mathbb{L}^2(w)) \bigcap \mathbb{L}^2((0,T),\mathbb{H}^1(w))$;*
- *for any function $\psi \in \mathbb{H}^1(w)$, we have :*

$$\partial_t \int_{\mathcal{D}} u_t\psi w = -\int_{\mathcal{D}} u_t \nabla V\cdot(w\nabla\psi + \psi\nabla w) - \int_{\mathcal{D}} \nabla u_t\cdot(w\nabla\psi + \psi\nabla w) + \int_{\mathcal{D}} (\partial_1\psi) u_t \frac{\varphi_\eta * u_t^{\partial_1 V}}{\varphi_\eta * u_t^1} w \tag{5.35}$$

  in the sense of distributions in time ;
- *$u_0 = p_0$.*

As for Definition 5.2.1, these conditions make sense.

With this definition, one has the following result :

Lemma 5.3.7 *Consider both the compact and the non compact cases. Under Assumptions i and ii, the distribution P_t^η of $\{X_t^\eta\}$ has a density p_t^η with respect to the Lebesgue measure such that p^η satisfies* (5.33) *in the sense of Definition 5.3.6.*

Moreover, the family $(p^\eta)_{\eta>0}$ is bounded in $\mathbb{L}^\infty((0,T),\mathbb{L}^2(\mathcal{D})) \cap \mathbb{L}^2((0,T),\mathbb{H}^1(\mathcal{D}))$ and, in the non compact case, under Assumption iii, $(p^\eta)_{\eta>0}$ is bounded in $\mathbb{L}^\infty((0,T),\mathbb{L}^2(w)) \cap \mathbb{L}^2((0,T),\mathbb{H}^1(w))$.

Proof. Since the drift coefficient in (5.28) is bounded, following the proof of Lemmas 5.2.3 and 5.2.4, we obtain that P_t^η has a density p_t^η, where p^η satisfies the first condition in Definition 5.3.6. Applying Itō's formula to $\psi(X_t^\eta)$ for some smooth ψ, we find that (5.34) ((5.35) in the non compact case) holds for a smooth ψ. Using the density of smooth functions in $\mathbb{H}^1(\mathbb{T}\times\mathbb{R}^{d-1})$, it holds for any ψ in $\mathbb{H}^1(\mathbb{T}^d)$, and the same is true for $\mathbb{H}^1(w)$ in the non compact case.

To prove that p^η is bounded independently of η, notice that from the boundedness of ∇V, the function $\frac{\varphi_\eta * p_t^{\eta,\partial_1 V}}{\varphi_\eta * p_t^{\eta,1}}$ is bounded from above uniformly with respect to η. Consequently, from Cauchy-Schwarz inequality,

$$\begin{aligned}\frac{1}{2}\partial_t \|p_t^\eta\|^2_{\mathbb{L}^2(\mathbb{T}\times\mathbb{R}^{d-1})} &= -\|\nabla p_t^\eta\|^2_{\mathbb{L}^2(\mathbb{T}\times\mathbb{R}^{d-1})} - \int_{\mathbb{T}\times\mathbb{R}^{d-1}} p_t^\eta \nabla p_t^\eta \cdot \nabla V + \int_{\mathbb{T}\times\mathbb{R}^{d-1}} (\partial_1 p_t^\eta) p_t^\eta \frac{\varphi_\eta * p_t^{\eta,\partial_1 V}}{\varphi_\eta * p_t^{\eta,1}} \\ &\leq -\|\nabla p_t^\eta\|^2_{\mathbb{L}^2(\mathbb{T}\times\mathbb{R}^{d-1})} + K\|p_t^\eta\|_{\mathbb{L}^2(\mathbb{T}\times\mathbb{R}^{d-1})}\|\nabla p_t^\eta\|_{\mathbb{L}^2(\mathbb{T}\times\mathbb{R}^{d-1})}.\end{aligned}$$

where, the constant K does not depend on η. We finish the proof using Young's inequality, and then Grönwall's Lemma.

The proof is similar in the non compact case.

Thanks to Lemma 5.3.7, we can prove the relative compactness of the family p^η in a nice sense.

Lemma 5.3.8 *Consider both the compact and the non compact cases. Under Assumptions i and ii, for any bounded open domain $\mathcal{O}$ in $\mathcal{D}$, the set $(p^\eta{}_{|\mathcal{O}})_{\eta>0}$ of restrictions of the functions p^η to $\mathcal{O}$ is relatively compact in the space $\mathbb{L}^2((0,T),\mathbb{L}^2(\mathcal{O}))$. Moreover, the set $(P^\eta)_{\eta>0}$ of laws of the solution is tight.*

Proof. We first prove the relative compactness of p^η in $\mathbb{L}^2((0,T),\mathbb{L}^2(\mathcal{O}))$. We use the fact that for a bounded open domain $\mathcal{O}$ and for $p,q\in(1,\infty)$, the space

$$E_{p,q} = \{f \in \mathbb{L}^p((0,T),\mathbb{H}^1(\mathcal{O})),\ \text{such that } \partial_t f \in \mathbb{L}^q((0,T),\mathbb{H}^{-1}(\mathcal{O}))\}$$

imbeds compactly in $\mathbb{L}^p((0,T),\mathbb{L}^2(\mathcal{O}))$ (see [50, page 57]). We already know that the set $(p^\eta)_{\eta>0}$ is bounded in $\mathbb{L}^2((0,T),\mathbb{H}^1(\mathcal{D}))$, so that the set $(p^\eta{}_{|\mathcal{O}})_{\eta>0}$ is bounded in $\mathbb{L}^2((0,T),\mathbb{H}^1(\mathcal{O}))$. Thus, it is enough to show that $(\partial_t p^\eta{}_{|\mathcal{O}})_{\eta>0}$ is bounded in $\mathbb{L}^q((0,T),\mathbb{H}^{-1}(\mathcal{O}))$, for some $q\in(1,\infty)$ to finish the proof. The following equation holds

$$\partial_t p^\eta = \operatorname{div}(p^\eta \nabla V) + \Delta p^\eta - \partial_1\left(p^\eta \frac{\varphi_\eta * p_t^{\eta,\partial_1 V}}{\varphi_\eta * p_t^{\eta,1}}\right),$$

showing, since $(p^\eta)_{\eta>0}$ is bounded in $\mathbb{L}^2((0,T),\mathbb{H}^1(\mathcal{D}))$, that $(\partial_t p^\eta)_{\eta>0}$ is bounded in $\mathbb{L}^2((0,T),\mathbb{H}^{-1}(\mathcal{D}))$, thus, $\partial_t p^\eta{}_{|\mathcal{O}}$ is bounded in $\mathbb{L}^2((0,T),\mathbb{H}^{-1}(\mathcal{O}))$. This shows that $(p^\eta_{|\mathcal{O}})_{\eta>0}$ is relatively compact in $\mathbb{L}^2((0,T),\mathbb{L}^2(\mathcal{O}))$.

Now we prove the relative compactness of $(P^\eta)_{\eta>0}$ in $\mathcal{P}(\mathcal{C}_T)$. For this aim, we use Kolmogorov compactness criterion. At time $t=0$, X_0^η is equal to X_0, independently of η. Consequently, the family $(X_0^\eta)_{\eta>0}$ is tight. To conclude the proof, it is enough to show that for some positive constants a, b and K,

$$\sup_{\eta>0}\mathbb{E}\left[|X_t^\eta - X_s^\eta|^a\right] \leq K|t-s|^{1+b}$$

for $0\leq s,t\leq T$. Since ∇V is bounded, we have, for $0\leq s,t\leq T$ and $p>1$,

$$\begin{aligned}\mathbb{E}\left[|X_t^\eta - X_s^\eta|^p\right]^{1/p} &\leq \mathbb{E}\left[\left|\int_s^t \nabla V(X_\tau^\eta)\mathrm{d}\tau\right|^p\right]^{1/p} + \mathbb{E}\left[|W_t - W_s|^p\right]^{1/p} + \mathbb{E}\left[\left|\int_s^t \frac{\varphi_\eta * u_\tau^{\eta,\partial_1 V}(X_\tau^{\eta,1})}{\varphi_\eta * u_\tau^{\eta,1}(X_\tau^{\eta,1})}\mathrm{d}\tau\right|^p\right]^{1/p} \\ &\leq K\left(|t-s| + |t-s|^{1/2}\right).\end{aligned}$$

This rewrites

$$\mathbb{E}\left[|X_t^\eta - X_s^\eta|^p\right] \leq K|t-s|^{p/2},$$

for some positive K. Taking $p=3$, Lemma 5.3.8 follows.

As a consequence of Lemma 5.3.8, using a diagonal argument, we can extract a subsequence of $\eta \to 0$, still denoted η such that :

- p^η converges almost everywhere on $(0,T)\times\mathcal{D}$ and in $\mathbb{L}^2((0,T),\mathbb{L}^2(\mathcal{O})) = \mathbb{L}^2((0,T)\times\mathcal{O})$ as η goes to 0, for any bounded open domain $\mathcal{O}$ to a function p^0.
- P^η converges in $\mathcal{P}(\mathcal{C}([0,T]))$ as η goes to 0 to a probability measure P^0.

To let η go to zero in (5.28), we finally need the following lemma.

Lemma 5.3.9 *Consider both the compact and the non compact cases. Under Assumptions i and ii, the limit p^0 of p^η is such that p^0_t is the density of the time marginal of P^0 for almost all times t. Moreover, the convergence of p^η to p^0 also holds in $\mathbb{L}^1((0,T)\times\mathcal{D})$ and up to a second subsequence extraction, $\frac{\varphi_\eta * p^{\eta,\partial_1 V}}{\varphi_\eta * p^{\eta,1}}$ converges almost everywhere on $(0,T)\times\mathbb{T}$ to $\frac{p^{0,\partial_1 V}}{p^{0,1}}$ as η goes to zero.*

Proof. We first prove that p^η converges to p^0 in $\mathbb{L}^1((0,T)\times\mathcal{D})$. One has

$$\begin{aligned}\int_0^T\int_{\mathcal{D}} |p^\eta - p^0| &= \int_0^T\int_{\mathcal{D}} (p^\eta - p^0) + 2\int_0^T\int_{\mathcal{D}} (p^\eta - p^0)^- \\ &= 2\int_0^T\int_{\mathcal{D}} (p^\eta - p^0)^-.\end{aligned}$$

But p^η converges almost everywhere to p^0, and $(p^\eta - p^0)^-$ is bounded from above by the integrable function p^0. Consequently, by the Lebesgue theorem, p^η converges to p^0 in $\mathbb{L}^1((0,T)\times\mathcal{D})$.

A consequence of this convergence and of the boundedness of $\partial_1 V$ is that the sequences $(p^{\eta,\partial_1 V})_{\eta>0}$ and $(p^{\eta,1})_{\eta>0}$ converge in $\mathbb{L}^1((0,T)\times\mathbb{T})$ respectively to $p^{0,\partial_1 V}$ and $p^{0,1}$.

As a consequence, $\varphi_\eta * p^{\eta,1}$ and $\varphi_\eta * p^{\eta,\partial_1 V}$ also converge in $\mathbb{L}^1((0,T)\times\mathbb{T})$ to the same limits. Therefore, up to the extraction of a second subsequence, we have pointwise convergence almost everywhere for the denominator and the numerator of $\frac{\varphi_\eta * p^{\eta,\partial_1 V}}{\varphi_\eta * p^{\eta,1}}$.

Now we show that p^0_t is for almost all t the density of the time marginal P^0_t of P^0. Since P^η converges to P^0 as η goes to 0 in $\mathcal{P}(\mathcal{C}_T)$, then $\mathbb{E}\left[\Psi(X^\eta)\right]$ converges to $\mathbb{E}\left[\Psi(X^0)\right]$ as η goes to 0, for any bounded continuous functional Ψ on $\mathcal{C}_T$. Taking a function of the form $\Psi(Y) = \int_0^T \theta(t)\tilde{\Psi}(Y_t)\mathrm{d}t$ where $\tilde{\Psi}$ and θ are bounded and continuous, one has

$$\mathbb{E}[\Psi(X^\eta)] = \int_0^T \theta(t)\left(\int_{\mathbb{T}\times\mathbb{R}^{d-1}} \tilde{\Psi} p^\eta_t\right)\mathrm{d}t.$$

Moreover, since p^η converges to p^0 in $\mathbb{L}^1((0,T)\times\mathbb{T}\times\mathbb{R}^{d-1})$, one has

$$\int_0^T \left(\theta(t)\int_{\mathbb{T}\times\mathbb{R}^{d-1}} \tilde{\Psi} p^\eta_t\right)\mathrm{d}t \to_{\eta\to 0} \int_0^T \theta(t)\left(\int_{\mathbb{T}\times\mathbb{R}^{d-1}} \tilde{\Psi} p^0_t\right)\mathrm{d}t.$$

As a result,

$$\mathbb{E}\left[\int_0^T \theta(t)\tilde{\Psi}(X^0_t)\mathrm{d}t\right] = \int_0^T \theta(t)\left(\int_{\mathbb{T}\times\mathbb{R}^{d-1}} \tilde{\Psi} p^0_t\right)\mathrm{d}t,$$

so that, almost everywhere, p^0_t is the time marginal of P^0.

We can now prove Theorem 5.3.5. We want to prove that P^0 is a solution to the martingale problem defined in Definition 5.3.3. It is enough to show that for $0 \le s_1 \le \ldots \le s_n \le s \le t$, any bounded continuous function g and any twice differentiable function ψ with bounded derivatives, one has $\int_{\mathcal{C}_T} g(x_{s_1},\ldots,x_{s_n})(m_t - m_s)\mathrm{d}P^0 = 0$.

Under the probability measure P^η, the canonical process $x \in \mathcal{C}([0,T])$ is such that

$$m_t^\eta = \psi(x_t) - \psi(x_0) - \int_0^t \Delta\psi(x_s)\mathrm{d}s + \int_0^t \nabla V(x_s)\nabla\psi(x_s)\mathrm{d}s - \int_0^t \partial_1\psi(x_s)\frac{\varphi_\eta * p_s^{\eta,\partial_1 V}(x_s^1)}{\varphi_\eta * p_s^{\eta,1}(x_s^1)}\mathrm{d}s$$

is a martingale. We thus have

$$\int_{\mathcal{C}_T} g(x_{s_1},\dots,x_{s_n})(m_t^\eta - m_s^\eta)\mathrm{d}P^\eta = 0.$$

Consequently, denoting $\tilde\eta = (\tilde\varepsilon, \tilde\alpha)$

$$\begin{aligned}\left|\int_{\mathcal{C}_T} g(x_{s_1},\dots,x_{s_n})(m_t - m_s)\mathrm{d}P^0\right| \le& \left|\int_{\mathcal{C}_T} g(x_{s_1},\dots,x_{s_n})(m_t^{\tilde\eta} - m_s^{\tilde\eta})\mathrm{d}P^\eta\right| \\ &+ \left|\int_{\mathcal{C}_T} g(x_{s_1},\dots,x_{s_n})(m_t^{\tilde\eta} - m_s^{\tilde\eta})\mathrm{d}(P^\eta - P^0)\right| \\ &+ \left|\int_{\mathcal{C}_T} g(x_{s_1},\dots,x_{s_n})\big((m_t^{\tilde\eta} - m_s^{\tilde\eta}) - (m_t - m_s)\big)\mathrm{d}P^0\right|.\end{aligned}$$

Taking $\limsup_{\tilde\eta\to 0}\limsup_{\eta\to 0}$, we obtain :

$$\left|\int_{\mathcal{C}_T} g(x_{s_1},\dots,x_{s_n})(m_t - m_s)\mathrm{d}P^0\right| \le \limsup_{\tilde\eta\to 0}\limsup_{\eta\to 0}\left|\int_{\mathcal{C}_T} g(x_{s_1},\dots,x_{s_n})(m_t^{\tilde\eta} - m_s^{\tilde\eta})\mathrm{d}P^\eta\right|. \tag{5.36}$$

Indeed, $g(x_{s_1},\dots,x_{s_n})(m_t^{\tilde\eta} - m_s^{\tilde\eta})$ is a bounded continuous function of x, and P^η converges to P^0. Moreover, we have

$$\begin{aligned}&\left|\int_{\mathcal{C}_T} g(x_{s_1},\dots,x_{s_n})\big((m_t^{\tilde\eta} - m_s^{\tilde\eta}) - (m_t - m_s)\big)\mathrm{d}P^0\right| \\ =&\left|\int_{\mathcal{C}_T}\int_s^t g(x_{s_1},\dots,x_{s_2})\partial_1\psi(x_\tau)\left[\frac{\varphi_{\tilde\eta} * p_\tau^{\tilde\eta,\partial_1 V}}{\varphi_{\tilde\eta} * p_\tau^{\tilde\eta,1}} - \frac{p_\tau^{0,\partial_1 V}}{p_\tau^{0,1}}\right](x_\tau^1)\ \mathrm{d}\tau\ \mathrm{d}P^0(x)\right| \\ \le& K\int_s^t\int_{\mathcal{D}}\left|\left[\frac{\varphi_{\tilde\eta} * p_\tau^{\tilde\eta,\partial_1 V}}{\varphi_{\tilde\eta} * p_\tau^{\tilde\eta,1}} - \frac{p_\tau^{0,\partial_1 V}}{p_\tau^{0,1}}\right](y)\right| p_\tau^0(y)\ \mathrm{d}y\ \mathrm{d}\tau.\end{aligned}$$

This last integral goes to 0 as $\tilde\eta$ goes to 0, since the function $\left[\frac{\varphi_\eta * p_\tau^{\tilde\eta,\partial_1 V}}{\varphi_{\tilde\eta} * p_\tau^{\tilde\eta,1}} - \frac{p_\tau^{0,\partial_1 V}}{p_\tau^{0,1}}\right]$ converges almost everywhere to 0 on $[s,t]\times\mathcal{D}$, and is bounded fom above by some positive constant. To conclude, we estimate the right hand side in (5.36) :

$$\begin{aligned}&\left|\int_{\mathcal{C}_T} g(x_{s_1},\dots,x_{s_n})(m_t^{\tilde\eta} - m_s^{\tilde\eta})\mathrm{d}P^\eta(x)\right| \\ =&\left|\int_{\mathcal{C}_T} g(x_{s_1},\dots,x_{s_n})((m_t^{\tilde\eta} - m_s^{\tilde\eta}) - (m_t^\eta - m_s^\eta))\mathrm{d}P^\eta(x)\right| \\ =&\left|\int_{\mathcal{C}_T} g(x_{s_1},\dots,x_{s_n})\int_s^t \partial_1\psi(x_\tau)\left(\frac{\varphi_\eta * p_\tau^{\eta,\partial_1 V}}{\varphi_\eta * p_\tau^{\eta,1}} - \frac{\varphi_{\tilde\eta} * p_\tau^{\tilde\eta,\partial_1 V}}{\varphi_{\tilde\eta} * p_\tau^{\tilde\eta,1}}\right)(x_\tau^1)\mathrm{d}\tau\mathrm{d}P^\eta(x)\right| \\ \le& K\int_s^t\int_{\mathcal{C}_T}\left|\frac{\varphi_\eta * p_\tau^{\eta,\partial_1 V}(x_\tau^1)}{\varphi_\eta * p_\tau^{\eta,1}(x_\tau^1)} - \frac{\varphi_{\tilde\eta} * p_\tau^{\tilde\eta,\partial_1 V}(x_\tau^1)}{\varphi_{\tilde\eta} * p_\tau^{\tilde\eta,1}(x_\tau^1)}\right|\mathrm{d}P^\eta(x)\mathrm{d}\tau \\ =& K\int_s^t\int_{\mathbb{T}}\left|\frac{\varphi_\eta * p_\tau^{\eta,\partial_1 V}(y)}{\varphi_\eta * p_\tau^{\eta,1}(y)} - \frac{\varphi_{\tilde\eta} * p_\tau^{\tilde\eta,\partial_1 V}(y)}{\varphi_{\tilde\eta} * p_\tau^{\tilde\eta,1}(y)}\right| p_\tau^{\eta,1}(y)\mathrm{d}y\mathrm{d}\tau.\end{aligned}$$

This last integral tends to 0 as η and $\tilde{\eta}$ go to 0, since $p^{\eta,1}$ converges in $\mathbb{L}^1((s,t)\times\mathbb{T})$, and since $\frac{\varphi_\eta * p_\tau^{\eta,\partial_1 V}(y)}{\varphi_\eta * p_\tau^{\eta,1}(y)} - \frac{\varphi_{\tilde{\eta}} * p_\tau^{\tilde{\eta},\partial_1 V}(y)}{\varphi_{\tilde{\eta}} * p_\tau^{\tilde{\eta},1}(y)}$ is bounded and converges almost everywhere to 0. We then obtain Theorem 5.3.5.

5.3.3 Another existence result for the nonlinear process

From Theorem 5.3.5, we know that existence holds for (5.6) under some regularity assumptions on the initial condition. Indeed, if P^0 is the limit of some subsequence of P^η, then the canonical process x defined on the canonical space $(\mathcal{C}_T, P^0)$ is a solution to Equation (5.6). By approximating the initial condition by regular densities, one can relax the regularity assumption.

Theorem 5.3.10 *Consider both the compact and non compact cases. Under Assumption i, weak existence holds for Equation* (5.6) *with given initial condition* X_0*. Moreover, for positive* s*, the law of* $\{X_s\}$ *has a density* p_s *such that, for* $0 < t < T$*,*

$$p \in \mathbb{L}^\infty((t,T),\mathbb{L}^2(\mathcal{D}))\bigcap\mathbb{L}^2((t,T),\mathbb{H}^1(\mathcal{D})).$$

Notice that, under the hypotheses of Theorem 5.3.10, we have no uniqueness result.

Proof. Theorem 5.3.5 yields existence for (5.6) when the initial condition satisfies Assumption ii. To prove existence without assumption on the initial condition, we use approximations of the initial condition. Let $(p_0^k)_{k\in\mathbb{N}}$ be a sequence of probability densities satisfying Assumption ii and converging to p^0 in $\mathcal{P}(\mathcal{D})$ (for example, this sequence can be obtained by convolution with a gaussian kernel). From Theorem 5.3.5, there exists a solution (X_t^k) to Equation (5.6) driven by a Brownian motion W defined on some probability space $(\Omega,\mathcal{F},\mathbb{P})$, such that X_0^k admits p_0^k as density.

As in the proof of Lemma 5.3.8, we can apply Kolmogorov criterion, so that the family of distributions P^k of $(\{X_t^k\})_{0\le t\le T}$ is tight. Consequently, we can extract from (P^k) a converging subsequence whose limit is denoted P. To prove that P satisfies the martingale problem defined in Definition 5.3.3, we need some estimate on the time marginals of P^k, uniformly in k.

According to Lemma 5.2.4, the law of $\{X_t^k\}$ has a density p_t^k such that p^k lies in $\mathbb{L}^\infty((0,T),\mathbb{L}^2(\mathcal{D}))$ and $\mathbb{L}^2((0,T),\mathbb{H}^1(\mathcal{D}))$. Notice that the drift coefficient $b_t^k(X_t) \stackrel{\text{def}}{=} -\nabla V(X_t) + \mathbb{E}[\partial_1 V(X_t)|\{X_t^1\}]e_1$ in Equation (5.6) is bounded, so that we can apply the Girsanov Theorem. Indeed, define

$$L_t^k = \exp\left(-\frac{1}{\sqrt{2}}\int_0^t b_s^k(X_s^k)\mathrm{d}W_s - \frac{1}{4}\int_0^t \|b_s^k(X_s^k)\|^2\mathrm{d}s\right).$$

Novikov's Condition is satisfied for this process, so that the formula

$$\mathbb{Q}_k(A) = \mathbb{E}[\mathbf{1}_A L_t^k],$$

for $A\in\sigma(W_s)_{s\le t}$, defines a probability distribution $\mathbb{Q}_k$ on Ω such that, under $\mathbb{Q}_k$, the process

$$\frac{1}{\sqrt{2}}\left(X_t^k - X_0^k\right) = W_t + \frac{1}{\sqrt{2}}\int_0^t b_s^k(X_s^k)\mathrm{d}s$$

is a Brownian motion. Denote γ_t^k the law of $\{X_t^k\}$ under $\mathbb{Q}_k$. Notice that since, under $\mathbb{Q}_k$, X_t^k is the sum of $\sqrt{2}$ times a Brownian motion at time t and an independent random variable X_0^k, γ_t^k has a density with respect to the Lebesgue measure which is bounded by $\frac{K}{t^{d/2}}$ where K is a constant not depending on k and t. As a result, for a given function ψ in $\mathbb{L}^2(\mathcal{D})$, one has

$$\begin{aligned}\left|\int_{\mathcal{D}}\psi(x)\mathrm{d}P_t^k(x)\right| &= \left|\mathbb{E}\left[\psi\left(X_t^k\right)\right]\right| \\ &= \left|\mathbb{E}_{\mathbb{Q}_k}\left[\psi\left(X_t^k\right)\left(L_t^k\right)^{-1}\right]\right| \\ &\leq \left(\int_{\mathcal{D}}\psi^2\mathrm{d}\gamma_t^k\right)^{\frac{1}{2}}\mathbb{E}\left[(L_t^k)^{-2}\right]^{1/2} \\ &\leq \frac{K}{t^{d/4}}\left\|\psi\right\|_{\mathbb{L}^2(\mathcal{D})},\end{aligned}$$

where K is a positive constant, which does not depend on k since $\left|b^k\right|$ is bounded from above by $\|\nabla V\|_{\mathbb{L}^\infty}$. Consequently, for any $0 < t < T$, $\|p_s^k\|_{\mathbb{L}^2(\mathcal{D})}$ is bounded uniformly in k and in $s \in [t,T]$. Moreover, since p^k is a solution to Equation (5.12) in the sense of Definition 5.2.1 one has, from (5.13)

$$\partial_t\|p_s^k\|^2_{\mathbb{L}^2(\mathcal{D})} \leq -\|\nabla p_s^k\|^2_{\mathbb{L}^2(\mathcal{D})} + K\|p_s^k\|^2_{\mathbb{L}^2(\mathcal{D})},$$

so that $(p^k)_{k\in\mathbb{N}}$ is also bounded in $\mathbb{L}^2((t,T),\mathbb{H}^1(\mathcal{D}))$. Adapting the proof of Lemma 5.3.8, we find that the family $(p^k_{|\mathcal{O}})$ is compact in $\mathbb{L}^2((t,T),\mathbb{L}^2(\mathcal{O}))$ for any open subset $\mathcal{O}$ of $\mathcal{D}$. By a diagonal argument, and using the proof of Lemma 5.3.9 we can thus construct a subsequence k_n such that

- P^{k_n} converges to a probability measure P^0 whose time marginals P_t have a density p_t^0, for all $t > 0$,
- p^{k_n} converges almost everywhere on $(0,T)\times\mathcal{D}$ and in $\mathbb{L}^1((0,T)\times\mathcal{D})$ to p^0,
- $\dfrac{p^{k_n,\partial_1 V}}{p^{k_n,1}}$ converges almost everywhere on $(0,T)\times\mathcal{D}$ to $\dfrac{p^{0,\partial_1 V}}{p^{0,1}}$.

Then, adapting the proof of Theorem 5.3.5, we see that P^0 solves the martingale problem.

5.3.4 Rate of convergence

We are going to exhibit a control on the rate of the convergence of p^η to p. Moreover, we give an estimate of the difference between $\dfrac{\varphi_\eta * p^{\eta,\partial_1 V}}{\varphi_\eta * p^{\eta,1}}$ and the biasing force $A'_t = \dfrac{p_t^{\partial_1 V}}{p_t^1}$ which is the quantity one is interested in practice.

Theorem 5.3.11 *Under Assumptions i and ii, one has, in the compact case,*

$$\|p^\eta - p\|_{\mathbb{L}^\infty((0,T),\mathbb{L}^2(\mathbb{T}^d))} + \|p^\eta - p\|_{\mathbb{L}^2((0,T),\mathbb{H}^1(\mathbb{T}^d))} \leq K(\alpha + \sqrt{\varepsilon}),$$

and, in the non compact case, under the additional Assumption iii

$$\|p^\eta - p\|_{\mathbb{L}^\infty((0,T),\mathbb{L}^2(w))} + \|p^\eta - p\|_{\mathbb{L}^2((0,T),\mathbb{H}^1(w))} \leq K(\alpha + \sqrt{\varepsilon}),$$

for some positive constant K not depending on α and ε. Moreover, we have the following bound on the estimation of the biasing force :

$$\left\|\frac{\varphi_\eta * p_t^{\eta,\partial_1 V}}{\varphi_\eta * p_t^{\eta,1}} - \frac{p_t^{\partial_1 V}}{p_t^1}\right\|_{\mathbb{L}^2((0,T),\mathbb{L}^\infty(\mathbb{T}))} \leq K\left(\alpha + \sqrt{\varepsilon}\right).$$

Proof. We give the proof in the non compact case, the one in the compact case being very similar. Similar calculations as in the proof of Theorem 5.2.6 yield :

$$\begin{aligned}\frac{1}{2}\partial_t\|p_t - p_t^\eta\|^2_{\mathbb{L}^2(w)} + \|\nabla p_t - \nabla p_t^\eta\|^2_{\mathbb{L}^2(w)} &\leq K\|p_t - p_t^\eta\|_{\mathbb{L}^2(w)}\|\nabla p_t - \nabla p_t^\eta\|_{\mathbb{L}^2(w)} + K\|p_t - p_t^\eta\|^2_{\mathbb{L}^2(w)} \\ &\quad + \|\nabla p_t^\eta - \nabla p_t\|_{\mathbb{L}^2(w)}\left\|p_t\frac{p_t^{\partial_1 V}}{p_t^1} - p_t^\eta\frac{p_t^{\eta,\partial_1 V} * \varphi_\eta}{p_t^{\eta,1} * \varphi_\eta}\right\|_{\mathbb{L}^2(w)}.\end{aligned}$$

We now estimate the last term.

$$\begin{aligned}
\left\| p_t \frac{p_t^{\partial_1 V}}{p_t^1} - p_t^\eta \frac{p_t^{\eta,\partial_1 V} * \varphi_\eta}{p_t^{\eta,1} * \varphi_\eta} \right\|_{\mathbb{L}^2(w)} &\leq \left\| \frac{p_t}{p_t^1} \left(p_t^{\partial_1 V} - \varphi_\eta * p_t^{\eta,\partial_1 V} \right) \right\|_{\mathbb{L}^2(w)} \\
&\quad + \left\| p_t \ \varphi_\eta * p_t^{\eta,\partial_1 V} \left(\frac{1}{p_t^1} - \frac{1}{\varphi_\eta * p_t^{\eta,1}} \right) \right\|_{\mathbb{L}^2(w)} \\
&\quad + \left\| (p_t - p_t^\eta) \frac{\varphi_\eta * p_t^{\eta,\partial_1 V}}{\varphi_\eta * p_t^{\eta,1}} \right\|_{\mathbb{L}^2(w)} \\
&\leq \|p_t\|_{\mathbb{L}^2(w)} \left\| \frac{1}{p_t^1} \left(p_t^{\partial_1 V} - \varphi_\eta * p_t^{\eta,\partial_1 V} \right) \right\|_{\mathbb{L}^\infty(\mathbb{T})} \\
&\quad + \|p_t\|_{\mathbb{L}^2(w)} \left\| \frac{\varphi_\eta * p_t^{\eta,\partial_1 V}}{p_t^1 (\varphi_\eta * p_t^{\eta,1})} \left(\varphi_\eta * p_t^{\eta,1} - p_t^1 \right) \right\|_{\mathbb{L}^\infty(\mathbb{T})} \\
&\quad + \left\| (p_t - p_t^\eta) \frac{\varphi_\eta * p_t^{\eta,\partial_1 V}}{\varphi_\eta * p_t^{\eta,1}} \right\|_{\mathbb{L}^2(w)} .
\end{aligned}$$

From Lemma 5.1.2, p_t^1 is bounded from below uniformly in time. Using this together with the facts that $\partial_1 V$ is bounded and $p \in \mathbb{L}^\infty((0,T), \mathbb{L}^2(w))$, one obtains

$$\left\| p_t \frac{p_t^{\partial_1 V}}{p_t^1} - p_t^\eta \frac{p_t^{\eta,\partial_1 V} * \varphi_\eta}{p_t^{\eta,1} * \varphi_\eta} \right\|_{\mathbb{L}^2(w)} \leq K \left(\|p_t^{\partial_1 V} - \varphi_\eta * p_t^{\eta,\partial_1 V}\|_{\mathbb{L}^\infty(\mathbb{T})} + \|p_t^1 - \varphi_\eta * p_t^{\eta,1}\|_{\mathbb{L}^\infty(\mathbb{T})} + \|p_t - p_t^\eta\|_{\mathbb{L}^2(w)} \right).$$

Consequently, we have to estimate $\|p_t^1 - \varphi_\eta * p_t^{\eta,1}\|_{\mathbb{L}^\infty(\mathbb{T})}$ and $\|p_t^{\partial_1 V} - \varphi_\eta * p_t^{\eta,\partial_1 V}\|_{\mathbb{L}^\infty(\mathbb{T})}$. One has, for $\gamma \in (0, 1/2)$,

$$\begin{aligned}
\|p_t^{\partial_1 V} - \varphi_\eta * p_t^{\eta,\partial_1 V}\|_{\mathbb{L}^\infty(\mathbb{T})} &\leq \|\varphi_\eta * (p_t^{\partial_1 V} - p_t^{\eta,\partial_1 V})\|_{\mathbb{L}^\infty(\mathbb{T})} + \|p_t^{\partial_1 V} - \varphi_\eta * p_t^{\partial_1 V}\|_{\mathbb{L}^\infty(\mathbb{T})} \\
&\leq K\alpha + \|p_t^{\partial_1 V} - p_t^{\eta,\partial_1 V}\|_{\mathbb{L}^\infty(\mathbb{T})} + \|p_t^{\partial_1 V} - \varphi_\eta * p_t^{\partial_1 V}\|_{\mathbb{L}^\infty(\mathbb{T})} \\
&\leq K \left(\alpha + \|p_t^{\partial_1 V} - p_t^{\eta,\partial_1 V}\|_{\mathbb{H}^1(\mathbb{T})}^{1/2+\gamma} \|p_t^{\partial_1 V} - p_t^{\eta,\partial_1 V}\|_{\mathbb{L}^2(\mathbb{T})}^{1/2-\gamma} \right) + \|p_t^{\partial_1 V} - \varphi_\eta * p_t^{\partial_1 V}\|_{\mathbb{L}^\infty(\mathbb{T})} \\
&\leq K \left(\alpha + \|p_t - p_t^\eta\|_{\mathbb{H}^1(w)}^{1/2+\gamma} \|p_t - p_t^\eta\|_{\mathbb{L}^2(\mathbb{T})}^{1/2-\gamma} \right) + \|p_t^{\partial_1 V} - \varphi_\eta * p_t^{\partial_1 V}\|_{\mathbb{L}^\infty(\mathbb{T})}.
\end{aligned}$$

Likewise, we have

$$\|p_t^1 - \varphi_\eta * p_t^{\eta,1}\|_{\mathbb{L}^\infty(\mathbb{T})} \leq K\alpha + K\|p_t - p_t^\eta\|_{\mathbb{H}^1(w)}^{1/2+\gamma} \|p_t - p_t^\eta\|_{\mathbb{L}^2(\mathbb{T})}^{1/2-\gamma} + \|p_t^1 - \varphi_\eta * p_t^1\|_{\mathbb{L}^\infty(\mathbb{T})}.$$

To conclude, notice that, in view of Lemma 5.0.1, $p_t^{\partial_1 V}$ lies in $\mathbb{H}^1(\mathbb{T})$. Thus $p_t^{\partial_1 V}$ is Hölder continuous with exponent $1/2$ and constant $C\|p_t^{\partial_1 V}\|_{\mathbb{H}^1(\mathbb{T})}$ (see [20, Corollaire IX.13]). Consequently, since $\psi_\varepsilon \equiv 0$ outside $[-\varepsilon, \varepsilon]$,

$$\begin{aligned}\left|p_t^{\partial_1 V}(x)-\varphi_\eta * p_t^{\partial_1 V}(x)\right| &= \left|\alpha\int_{\mathbb{T}} p_t^{\partial_1 V}(x)\mathrm{d}x+\int_{\mathbb{T}}\psi_\varepsilon(y)\big(p_t^{\partial_1 V}(x)-p_t^{\partial_1 V}(x-y)\big)\mathrm{d}y\right| \\ &\le K\left(\alpha+\|p_t^{\partial_1 V}\|_{\mathbb{H}^1(\mathbb{T})}\int_{\mathbb{T}}\psi_\varepsilon(y)\sqrt{y}\mathrm{d}y\right) \\ &\le K\left(\alpha+\|p_t^{\partial_1 V}\|_{\mathbb{H}^1(\mathbb{T})}\sqrt{\varepsilon}\int_{\mathbb{T}}\psi_\varepsilon(y)\mathrm{d}y\right) \\ &\le K\left(\alpha+\sqrt{\varepsilon}\|p_t\|_{\mathbb{H}^1(w)}\right). \end{aligned} \tag{5.37}$$

The same inequality holds for p^1

$$\left|p_t^1(x)-\varphi_\eta * p_t^1(x)\right| \le K\left(\alpha+\sqrt{\varepsilon}\|p_t\|_{\mathbb{H}^1(w)}\right).$$

Gathering all the previous inequalities, we obtain,

$$\begin{aligned}\frac{1}{2}\partial_t\|p_t-p_t^\eta\|^2_{\mathbb{L}^2(w)}+\|\nabla p_t-\nabla p_t^\eta\|^2_{\mathbb{L}^2(w)} &\le K\|p_t-p_t^\eta\|_{\mathbb{L}^2(w)}\|\nabla p_t-\nabla p_t^\eta\|_{\mathbb{L}^2(w)}+K\|p_t-p_t^\eta\|^2_{\mathbb{L}^2(w)} \\ &\quad + K\|p_t-p_t^\eta\|^{1/2+\gamma}_{\mathbb{H}^1(w)}\|p_t-p_t^\eta\|^{1/2-\gamma}_{\mathbb{L}^2(\mathbb{T})}\|\nabla p_t-\nabla p_t^\eta\|_{\mathbb{L}^2(w)} \\ &\quad + K\left(\alpha+\sqrt{\varepsilon}\|p_t\|_{\mathbb{H}^1(w)}\right)\|\nabla p_t-\nabla p_t^\eta\|_{\mathbb{L}^2(w)} \\ &\quad + K\alpha\|\nabla p_t-\nabla p_t^\eta\|_{\mathbb{L}^2(w)}.\end{aligned}$$

Consequently, from Young's inequality,

$$\partial_t\|p_t-p_t^\eta\|^2_{\mathbb{L}^2(w)}+\|\nabla p_t^\eta-\nabla p_t\|^2_{\mathbb{L}^2(w)} \le K\left(\|p_t-p_t^\eta\|^2_{\mathbb{L}^2(w)}+\alpha^2+\varepsilon\|p_t\|^2_{\mathbb{H}^1(w)}\right).$$

Grönwall's Lemma yields the first statement of Theorem 5.3.11, noticing that p lies in $\mathbb{L}^2((0,T),\mathbb{H}^1(w))$.

For the second statement, arguing as we did above, it holds that

$$\begin{aligned}\left\|\frac{\varphi_\eta * p_t^{\eta,\partial_1 V}}{\varphi_\eta * p_t^{\eta,1}}-\frac{p_t^{\partial_1 V}}{p_t^1}\right\|_{\mathbb{L}^\infty(\mathbb{T})} &\le K\left(\left\|p_t^1-\varphi_\eta * p_t^{\eta,1}\right\|_{\mathbb{L}^\infty(\mathbb{T})}+\left\|\varphi_\eta * p_t^{\eta,\partial_1 V}-p_t^{\partial_1 V}\right\|_{\mathbb{L}^\infty(\mathbb{T})}\right) \\ &\le K\left(\alpha+\sqrt{\varepsilon}\|p_t\|_{\mathbb{H}^1(\mathbb{T})}+\|p_t-p_t^\eta\|_{\mathbb{H}^1(w)}\right).\end{aligned}$$

We finish the proof by squaring this inequality and then integrating.

5.4 An interacting particle system approximation

In this section, we prove the convergence of the interacting particle system to the regularized nonlinear processes, and we estimate the difference between the regularized biasing force $\frac{\varphi_\eta * p_t^{\eta,\partial_1 V}}{\varphi_\eta * p_t^{\eta,1}}$ and its particle approximation.

Theorem 5.4.1 *Let T be a positive time. Under Assumptions i, ii, iv and v, the solution $(X^\eta_{t,n,N})_{N\ge 1}$ of (5.8) with initial condition $X^\eta_{0,n,N}=X_{0,n}$ converges to the solution $\bar{X}^\eta_{t,n}$ to (5.28) with initial condition $X_{0,n}$ in the following sense : for all $1\le n\le N$, and for $\varepsilon,\alpha\le 1$,*

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|\bar{X}^\eta_{t,n}-X^\eta_{t,n,N}\right|\right] < \frac{1}{\sqrt{N}}e^{\frac{K}{\alpha\varepsilon^2}},$$

K being some constant not depending on α, ε and N.

Moreover, one has

$$\mathbb{E}\left[\sup_{t\in[0,T],x^1\in\mathbb{T}}\left|\frac{\varphi_\eta * p^{\eta,\partial_1 V}}{\varphi_\eta * p^{\eta,1}}(x^1) - \frac{\sum_{n=1}^N \varphi_\eta(x^1 - X^{\eta,1}_{t,n,N})\partial_1 V(X^\eta_{t,n,N})}{\sum_{n=1}^N \varphi_\eta(x^1 - X^{\eta,1}_{t,n,N})}\right|\right] \leq \frac{1}{\sqrt{N}} e^{\frac{K}{\alpha\varepsilon^2}}. \tag{5.38}$$

Notice that the the right hand side of (5.38) explodes when ε goes to 0 for a fixed value of N, so that the size of ε has to be chosen carefully depending on the value of N. We will also investigate this point numerically in the next section.

To simplify notation, we omit the subscript N and the superscript η. We first establish the following inequality :

Lemma 5.4.2 *We have, for $\varepsilon, \alpha < 1$, and for any t in $(0,T]$,*

$$\left|X_{t,n} - \bar{X}_{t,n}\right| \leq \frac{K}{\alpha\varepsilon^2}\int_0^t \left(\left|X_{s,n} - \bar{X}_{s,n}\right| + \frac{1}{N}\sum_{m=1}^N \left|X_{s,m} - \bar{X}_{s,m}\right|\right)\mathrm{d}s + K\int_0^t A^{n,N}_s \mathrm{d}s,$$

where K does not depend on α, ε and t, and $A^{n,N}_t$ is defined by

$$A^{n,N}_t = \frac{1}{\alpha}\left(\left|\frac{1}{N}\sum_{m=1}^N \varphi_\eta(\bar{X}^1_{s,n} - \bar{X}^1_{s,m})\partial_1 V(\bar{X}_{s,m}) - \varphi_\eta * p^{\eta,\partial_1 V}_s(\bar{X}^1_{s,n})\right| + \left|\frac{1}{N}\sum_{m=1}^N \varphi_\eta(\bar{X}^1_{s,n} - \bar{X}^1_{s,m}) - \varphi_\eta * p^{\eta,1}_s(\bar{X}^1_{s,n})\right|\right).$$

Proof. From the definition of $X_{t,n}$ and $\bar{X}_{t,n}$, we have

$$\begin{aligned}\left|X_{t,n} - \bar{X}_{t,n}\right| \leq{}& \left|\int_0^t \left(\nabla V(X_{s,n}) - \nabla V(\bar{X}_{s,n})\right)\mathrm{d}s\right| \\ &+ \left|\int_0^t \frac{\sum_{m=1}^N \varphi_\eta(X^1_{s,n} - X^1_{s,m})\partial_1 V(X_{s,m})}{\sum_{m=1}^N \varphi_\eta(X^1_{s,n} - X^1_{s,m})}\mathrm{d}s - \int_0^t \frac{\varphi_\eta * p^{\eta,\partial_1 V}_s(\bar{X}^1_{s,n})}{\varphi_\eta * p^{\eta,1}_s(\bar{X}^1_{s,n})}\mathrm{d}s\right|.\end{aligned}$$

First, $\left|\int_0^t \left(\nabla V(X_{s,n}) - \nabla V(\bar{X}_{s,n})\right)\mathrm{d}s\right|$ is bounded from above by $K\int_0^t |X_{s,n} - \bar{X}_{s,n}|\mathrm{d}s$, since ∇V is Lipschitz continuous. Now, we decompose

$$\begin{aligned}&\left|\frac{\sum_{m=1}^N \varphi_\eta(X^1_{s,n} - X^1_{s,m})\partial_1 V(X_{s,m})}{\sum_{m=1}^N \varphi_\eta(X^1_{s,n} - X^1_{s,m})} - \frac{\varphi_\eta * p^{\eta,\partial_1 V}_s(\bar{X}^1_{s,n})}{\varphi_\eta * p^{\eta,1}_s(\bar{X}^1_{s,n})}\right| \\ &\leq \left|\frac{\sum_{m=1}^N \varphi_\eta(X^1_{s,n} - X^1_{s,m})\partial_1 V(X_{s,m})}{\sum_{m=1}^N \varphi_\eta(X^1_{s,n} - X^1_{s,m})} - \frac{\sum_{m=1}^N \varphi_\eta(\bar{X}^1_{s,n} - \bar{X}^1_{s,m})\partial_1 V(\bar{X}_{s,m})}{\sum_{m=1}^N \varphi_\eta(\bar{X}^1_{s,n} - \bar{X}^1_{s,m})}\right| \\ &\quad + \left|\frac{\sum_{m=1}^N \varphi_\eta(\bar{X}^1_{s,n} - \bar{X}^1_{s,m})\partial_1 V(\bar{X}_{s,m})}{\sum_{m=1}^N \varphi_\eta(\bar{X}^1_{s,n} - \bar{X}^1_{s,m})} - \frac{\varphi_\eta * p^{\eta,\partial_1 V}_s(\bar{X}^1_{s,n})}{\varphi_\eta * p^{\eta,1}_s(\bar{X}^1_{s,n})}\right|.\end{aligned} \tag{5.39}$$

Using Assumptions i and v, the first term in the right hand-side of (5.39) can be bounded by $\frac{K}{\alpha\varepsilon^2}\left(\left|X_{s,n} - \bar{X}_{s,n}\right| + \frac{1}{N}\sum_{m=1}^N |X_{s,m} - \bar{X}_{s,m}|\right)$, and the second term in the right hand side of (5.39) can be bounded by $KA^{n,N}_t$.

Proof (Proof of Theorem 5.4.1). As a consequence of Lemma 5.4.2, we get, for $\alpha, \varepsilon \leq 1$,

$$\sup_{t\in[0,T]}|X_{t,n} - \bar{X}_{t,n}| \leq \frac{K}{\alpha\varepsilon^2}\int_0^T \left(\sup_{s\in[0,t]}\left|X_{s,n} - \bar{X}_{s,n}\right| + \frac{1}{N}\sum_{m=1}^N \sup_{s\in[0,t]}|X_{s,m} - \bar{X}_{s,m}|\right)\mathrm{d}t + K\int_0^T A^{n,N}_t \mathrm{d}t.$$

Taking the expectation, and using the exchangeability of the couples $(X_n, \bar{X}_n)_{1\leq n\leq N}$, we get

$$\mathbb{E}\left[\sup_{t\in[0,T]}|X_{t,n}-\bar{X}_{t,n}|\right]\leq\frac{K}{\alpha\varepsilon^2}\int_0^T\mathbb{E}\left[\sup_{s\in[0,t]}|\bar{X}_{s,n}-X_{s,n}|\right]\mathrm{d}t+K\int_0^T\mathbb{E}\left[A_t^{n,N}\right]\mathrm{d}t.$$

By Grönwall's lemma, one obtains

$$\mathbb{E}\left[\sup_{t\in[0,T]}|X_{t,n}-\bar{X}_{t,n}|\right]\leq Ke^{\frac{K}{\alpha\varepsilon^2}T}\int_0^T\mathbb{E}\left[A_t^{n,N}\right]\mathrm{d}t.$$

To conclude, we estimate $\int_0^T\mathbb{E}[A_t^{n,N}]\mathrm{d}t$. Let

$$\Phi_t^m=\varphi_\eta(\bar{X}_{t,1}^1-\bar{X}_{t,m}^1)\partial_1V(\bar{X}_{t,m})-\varphi_\eta*p_t^{\eta,\partial_1V}(\bar{X}_{t,1}^1)$$

and

$$\Psi_t^m=\varphi_\eta(\bar{X}_{t,1}^1-\bar{X}_{t,m}^1)-\varphi_\eta*p_t^{\eta,1}(\bar{X}_{t,1}^1).$$

We have, for $t\leq T$, using again the exchangeability of the couples $(X_n,\bar{X}_n)_{1\leq n\leq N}$,

$$\begin{aligned}\left[\mathbb{E}A_t^{n,N}\right]^2&\leq\mathbb{E}\left[(A_t^{n,N})^2\right]\\&\leq\frac{K}{\alpha^2}\left(\mathbb{E}\left[\left(\frac{1}{N}\sum_{m=1}^N\Phi_t^m\right)^2\right]+\mathbb{E}\left[\left(\frac{1}{N}\sum_{m=1}^N\Psi_t^m\right)^2\right]\right)\\&=\frac{K}{N^2\alpha^2}\sum_{m,m'}\left(\mathbb{E}\left[\Phi_t^m\Phi_t^{m'}\right]+\mathbb{E}\left[\Psi_t^m\Psi_t^{m'}\right]\right).\end{aligned}$$

But the Φ_t^m and Ψ_t^m are centered for $m\geq2$, and, for $m\neq m'$, Φ_t^m and $\Phi_t^{m'}$, (as well as Ψ_t^m and $\Psi_t^{m'}$) are independent conditionally on $\bar{X}_{t,1}$. Therefore the double products vanish, and, by exchangeability

$$\left[\mathbb{E}A_t^{n,N}\right]^2\leq\frac{K(N-1)}{\alpha^2N^2}\left(\mathbb{E}\left[(\Phi_t^2)^2\right]+\mathbb{E}\left[(\Psi_t^2)^2\right]\right)+\frac{K}{\alpha^2N^2}(\mathbb{E}\left[(\Phi_t^1)^2\right]+\mathbb{E}\left[(\Psi_t^1)^2\right]).$$

But one has $\mathbb{E}\left[(\Phi_t^2)^2\right]+\mathbb{E}\left[(\Psi_t^2)^2\right]\leq K\varepsilon^{-2}$ and $\mathbb{E}\left[(\Phi_t^1)^2\right]+\mathbb{E}\left[(\Psi_t^1)^2\right]\leq K\varepsilon^{-2}$, and the first assertion in Theorem 5.4.1 follows.

For the estimation of the force, adapting the proof of Lemma 5.4.2, we see that

$$\begin{aligned}&\mathbb{E}\left[\sup_{t\in[0,T]}\left|\frac{\varphi_\eta*p_t^{\eta,\partial_1V}}{\varphi_\eta*p_t^{\eta,1}}(x^1)-\frac{\sum_{n=1}^N\varphi_\eta(x^1-X_{t,n,N}^1)\partial_1V(X_{t,n,N})}{\sum_{n=1}^N\varphi_\eta(x^1-X_{t,n,N}^1)}\right|\right]\\\leq&\frac{1}{\alpha}\mathbb{E}\left[\sup_{t\in[0,T]}\left|\frac{1}{N}\sum_{n=1}^N\varphi_\eta(x^1-\bar{X}_{t,n}^1)\partial_1V(\bar{X}_{t,n})-\varphi_\eta*p_t^{\eta,\partial_1V}(x^1)\right|\right]\\&+\frac{1}{\alpha}\mathbb{E}\left[\sup_{t\in[0,T]}\left|\frac{1}{N}\sum_{n=1}^N\varphi_\eta(x^1-\bar{X}_{t,n}^1)-\varphi_\eta*p_t^{\eta,1}(x^1)\right|\right]+\frac{K}{\alpha\varepsilon^2N}\mathbb{E}\left[\sup_{t\in[0,T]}\sum_{n=1}^N\left|X_{t,n}-\bar{X}_{t,n}\right|\right]\\\leq&\frac{1}{\sqrt{N}}e^{\frac{K}{\alpha\varepsilon^2}}.\end{aligned}$$

Indeed, $(\varphi_\eta(x^1-\bar{X}_{t,n}^1)\partial_1V(\bar{X}_{t,n})-\varphi_\eta*p_t^{\eta,\partial_1V}(x^1))_{n\in\{1\dots N\}}$, as well as $(\varphi_\eta(x^1-\bar{X}_{t,n}^1)-\varphi_\eta*p_t^{\eta,1}(x^1))_{n\in\{1\dots N\}}$, are i.i.d. centered random variables whose variance is bounded by $\frac{K}{\varepsilon^2}$, uniformly in time.

5.5 Numerical results

In this section we give some numerical simulations to illustrate our previous results. Here, the parameter α, which was introduced to enable theoretical estimations, is taken to be 0.

Notice that the discretization method used here in the simulations and based on a Nadaraya-Watson estimator is different from the classical implementations of the ABF technique [32]. Indeed, in molecular dynamics codes, time averages are used in order to smooth the problem : first the equation on A_t given in (5.4) is typically replaced by

$$\partial_t A'_t(z) = \frac{1}{\tau}(\mathbb{E}[F(X_t)|\xi(X_t) = z] - A_t)$$

which makes A_t vary more smoothly. Second, in addition to the particle approximation, an ergodic average for the computation of the conditional expectation in (5.4) is used.

In order to accelerate the convergence, one can also use a selection mechanism that gives more weight to particles located in less explored areas (see [47]).

5.5.1 Efficiency of the ABF method

Let us introduce a low dimensional example to illustrate the efficiency of the ABF method and its particle approximation.

In this first example, we simulated the particle approximation with 1000 particles, in the potential defined for (x, y) in $[-2, 2] \times \mathbb{R}$ by

$$V_1(x, y) = 5e^{-x^2-y^2} - 5e^{-(x-1)^2-y^2} - 5e^{-(x+1)^2-y^2} + 0.2x^4 + 0.2y^4, \tag{5.40}$$

and extended periodically in the x direction with period 4. The level sets of V_1 are depicted on Figure 5.1.

On Figure 5.1, we also plotted the position of the particles after 2000 iteration of an Euler-Maruyama approximation of Equation (5.8) with a time step of 0.01. The value of the parameters are $\varepsilon = 0.01$, $\beta = 10$ and $N = 1000$. On Figure 5.2, we plotted the graph of the mean force (computed by numerical integration, which is still possible due to the low dimensionnality), and the value of the approximate mean force computed on a regular grid. The $\mathbb{L}^1-$distance between the two functions is 6.93×10^{-2}, while the $\mathbb{L}^1-$norm of the function A' is 12.9.

Notice that without biasing force, one obtains a very poor sampling, since the particles do not escape from the well they started in : see Figure 5.3, where we plotted 200 independent simulations of a Langevin dynamics (5.1) using 2000 iterations of an Euler-Maruyama scheme of time step 0.01.

On Figure 5.4 we show the $\mathbb{L}^1$ distance between the actual value of the mean force A' and its approximation at time 20, obtained for one simulation of the system, as a function of the number of particles used in the simulation. Using a least square regression, we find that the slope of the curve is approximatively -0.59, which matches with the theoretical rate of $N^{-1/2}$.

5.5.2 Tuning of the parameters

In Theorem 5.4.1, we showed that the particle approximation converges as ε goes to 0 and N goes to infinity, provided that ε does not go to zero too fast compared to N. The practical difficulty that one encounters to apply this result is to choose a good scaling for ε in term of N.

On Figure 5.5, we can see the $\mathbb{L}^1$ error between the mean force and its aproximation at time 20, as a function of the parameter ε, using $N = 1500$ particles.

Actually, for a fixed value of N, there is only a small range of values for ε for which the error is small.

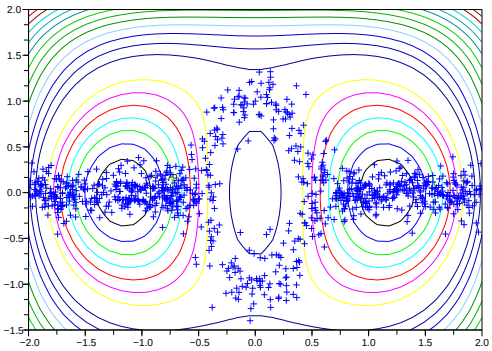

Fig. 5.1. Contour plot of the potential V_1 with the positions of 1000 particles at time $t = 20$.

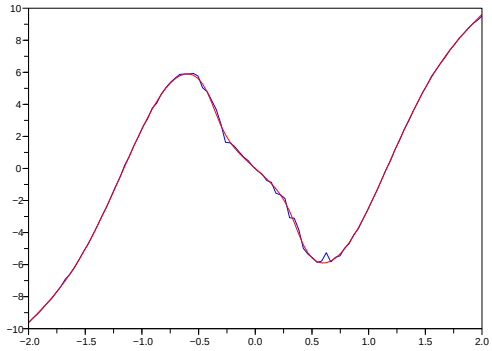

Fig. 5.2. Particle approximation of the mean force. The smooth curve is the actual value of the mean force, the rough one is the approximation.

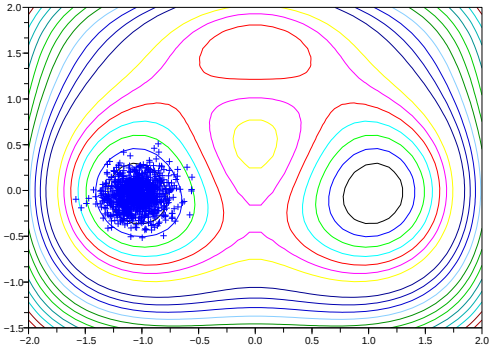

Fig. 5.3. 200 independent realizations of a Langevin dynamics at time $t = 20$.

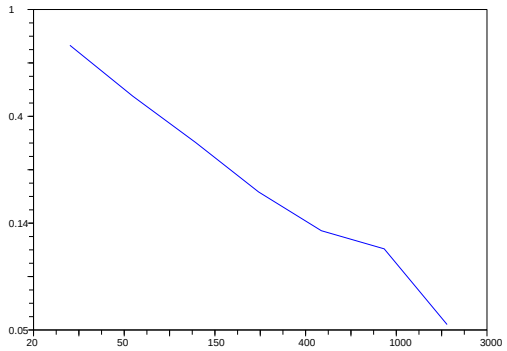

Fig. 5.4. Error as a function of N (logarithmic scale).

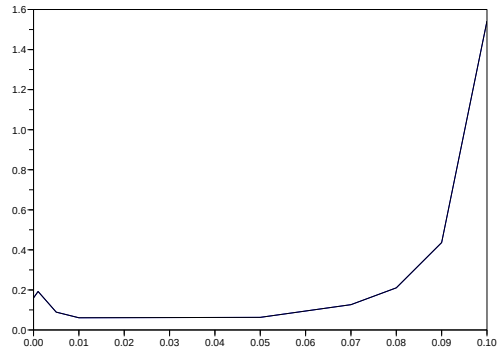

Fig. 5.5. Error as a function of ε.

First, the limit of the error as ε goes to 0 does not even vanish as N tends to infinity. The reason is that, since the particles interact with each other in a range of ε, the number of particles which interact with a given particle is of order εN. Hence, when ε tends to 0 while N is fixed, the particles cannot see each other. Therefore, the natural limit of the particle system in the limit $\varepsilon \to 0$, N fixed, should be a system of independent particles following the dynamics

$$\mathrm{d}X_t = (-\nabla V(X_t) + e_1 \partial_1 V(X_t))\,\mathrm{d}t + \sqrt{2\beta^{-1}}\mathrm{d}W_t.$$

Unfortunately, in the general case, the drift in the above dynamics is not obtained as the gradient of a potential, so that no invariant measure for X_t is known. This would consequently induce a non vanishing bias in the estimation of A.

For example, for the potential $V(x,y) = \frac{1}{2}(y - \sin(2\pi x))^2$, one can prove that the dynamics obtained by canceling the force on the reaction coordinate x, namely the couple $(\{X_t\}, Y_t)$ defined by the dynamics

$$\begin{cases} \mathrm{d}X_t &= \sqrt{2}\mathrm{d}W_t^1, \\ \mathrm{d}Y_t &= (-Y_t + \sin(2\pi X_t))\,\mathrm{d}t + \sqrt{2}\mathrm{d}W_t^2 \end{cases}$$

converges in law to the couple $\left(\xi, \int_0^\infty e^{-s}\sin(2\pi(\xi + \sqrt{2}W_s))\mathrm{d}s + G\right)$, where W is a standard Brownian motion, ξ is uniformly distributed on $\mathbb{T}$, and G is a standard normal random variable,

independent of W. This is not the correct limit distribution, since the law of Y conditionned to the value of $\{X\}$ should be Gaussian, which is not the case here.

For a large value of ε, the behavior of the particle system can be really different from the expected behavior of the dynamics (5.4). In the following example, the particles, instead of freely visiting the x axis, keep stuck in the local minima they started in. Indeed, the large value of ε made that the biasing term is close to the mean of $\partial_1 V(X^i)$ on all particles, whose value is close to 0. Consequently, the biasing force is not large enough to prevent the particle from being trapped in the local minima.

In the following example we considered the potential V_1 defined in (5.40), took $\varepsilon = 1$, and simulated 200 particles during 2000 iterations of time step 0.01. The result can be seen on Figure 5.6.

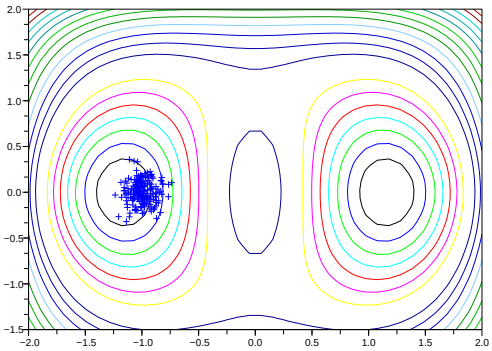

Fig. 5.6. Bad sampling due to a too large value of ε.

One way to increase the sample size while keeping the number N of particles fixed is to include time averages for the estimation of the conditional expectation. This is actually the common practice in the applied community (see [24, 32]).

5.5.3 Discussion on the choice of the reaction coordinate

We now give another example to illustrate the limitations of the ABF method. We consider the 4-periodical potential (in the x-direction) defined for (x, y) in $[-2, 2] \times \mathbb{R}$ by

$$V_2(x,y) = 3e^{-x^2-(y-1/3)^2} - 3e^{-x^2-(y-5/3)^2} - 5e^{-(x-1)^2-y^2} - 5e^{-(x+1)^2-y^2} + 0.2x^4 + 0.2(y-1/3)^4, \tag{5.41}$$

whose level sets are depicted on Figure 5.7. This potential has been introduced in [55].

The potential V_2 displays two deep minima approximately located at $(\pm 1, 0)$. There is a maximum located at $(0, 0.5)$, so that there are two possible paths between the main minima. The first one is a direct path meeting a saddle point approximately at $(0, -0.3)$. The other path goes through two saddle points at $(\pm 0.5, 1)$ and a small minima at $(0, 1.5)$. Even if the first path is more direct than the second one, the prefered path in low temperature regimes will be the second one, since its energy barrier is smaller.

We simulated the particle approximation of the ABF method with $N = 1000$ particles, window width $\varepsilon = 0.01$, after 2000 iterations of an Euler-Maruyama scheme of time step 0.01, and plotted the positions of the particles on Figure 5.7.

At the low temperature $\beta = 10$, the particles are expected to hop from one well to the other mainly through the upper channel, which is not the case here. This is due to a bad choice of

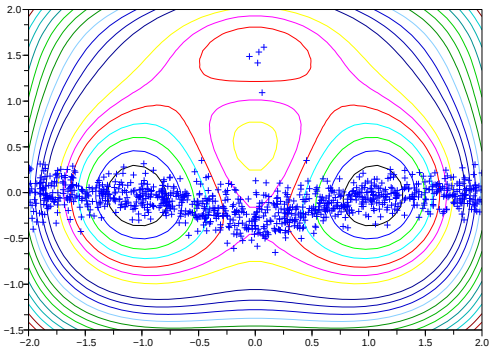

Fig. 5.7. Poor sampling due to a bad choice of the reaction coordinate.

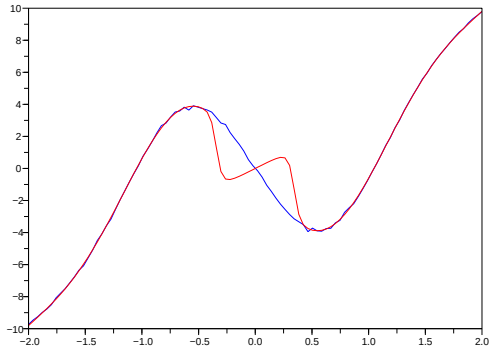

Fig. 5.8. Biased evaluation of the biasing force due to a bad choice of the reaction coordinate. The smooth curve is the value of the mean force. The rough curve is the approximation. Here, the approximation does not see the variations of the mean force around 0.

the reaction coordinate. Indeed, the biasing force only acts in the x direction, so that a particle trapped in the left side well will naturally escape through a horizontal path, and will take the lower channel. As a result, the computation of the force is clearly biased, because of the poor sampling of the upper channel, see Figure 5.8, the $\mathbb{L}^1-$distance between the two functions is of 0.4.

We still have convergence to the correct mean force, but at a slow rate, since the reaction coordinate has not been chosen in an optimal way. Indeed, with the same parameters, but after 2.10^6 iterations, the result is much better, see Figures 5.9 and 5.10. The $\mathbb{L}^1-$distance between the mean force and its approximation is of 0.15, while the function A' has $\mathbb{L}^1-$norm 10.9.

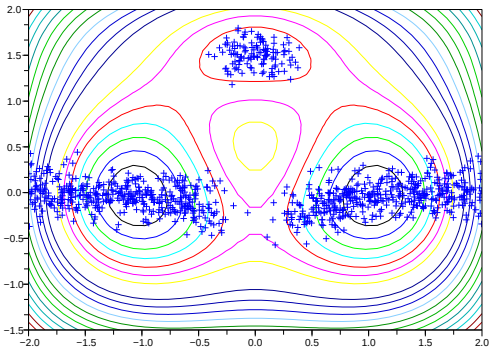

Fig. 5.9. Same simulation as on Figure 5.7 at time 2000.

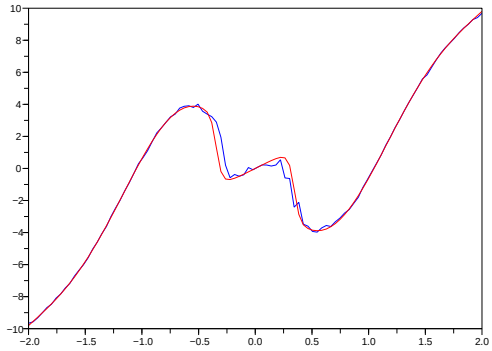

Fig. 5.10. Approximation of the free energy corresponding to Figure 5.9. The smooth curve is the free energy, the rough one is the approximation.

Partie III

Sensibilité d'une diffusion par rapport à un paramètre

6

Sensibilité d'une diffusion par rapport à un paramètre

Résumé : On présente des estimateurs d'ordre élevé utilisés en chimie quantique pour le calcul d'éléments propres d'opérateurs de Schrödinger. Précisément, on cherche à calculer l'énergie fondamentale de l'opérateur ainsi que sa dérivée par rapport à un paramètre. Dans le cadre de la méthode de Monte Carlo diffusive, le calcul de cet estimateur nécessite l'introduction d'un processus de diffusion dont le terme de dérive dépend d'un paramètre, ainsi que le calcul de la dérivée de ce processus par rapport à ce paramètre, appelée processus tangent. Nous montrons que la moyenne de l'estimateur peut être obtenue comme une moyenne en temps long sur les trajectoires de la diffusion et de son vecteur tangent. Un calcul efficace de cette moyenne en temps demande un contrôle sur la variance du processus tangent. Nous donnons quelques conditions assurant que cette variance est finie, et nous présentons quelques méthodes particulaires permettant de réduire cette variance.

Mots-Clés : Méthode de Monte Carlo variationnelle, méthode de Monte Carlo diffusive, équations différentielles stochastiques, réduction de variance, formules de Feynman-Kac.

Abstract : We present high-order estimators used in the fields of quantum Monte Carlo methods for the computation of eigenvalues of Schrödinger operators and of their derivatives with respect to some parameter. In order to be able to compute this high order estimator in the case of the Diffusion Monte Carlo method, we use a diffusion process whose drift term depends on some parameter. One then needs to compute the derivative of this process with respect to the parameter, called the tangent process. We show that the expectation of the estimator can be obtained as the long time average on the paths of the diffusion and its tangent vector. Efficient computation of the expectation requires a control of the variance of the tangent vector. We give conditions for this variance to be finite, and present some particle methods to reduce this variance.

Keywords : Variational Monte Carlo method, diffusion Monte Carlo method, stochastic differential equations, variance reduction, Feynman-Kac formulæ.

6.1 Eigenvalues of Schrödinger operators

In quantum mechanics, the state of a physical system is entirely described by a *wave function* which is a complex valued function whose squared modulus $|\varphi|^2$ is a probability density describing the probability of observation of the different states of the system, for a human observer. The fundamental equation of quantum physics is the following partial differential equation, the so-called *Schrödinger equation*, which describes the evolution of the wave function :

$$i\partial_t \varphi_t = \frac{1}{2}\Delta\varphi - V\varphi. \tag{6.1}$$

In equation (6.1), the function V is the potential energy associated to the system. Since $|\varphi|^2$ is a probability density, the natural function space for considering φ is the space $\mathbb{L}^2(\Omega)$ of square integrable functions, where Ω is the state space of the system, assumed to be some open connected subset of $\mathbb{R}^d$. The scalar product of $\mathbb{L}^2(\Omega)$ is denoted

$$\langle \varphi, \psi \rangle = \int_\Omega \varphi\psi.$$

For nice potentials, the Schrödinger operator $H\varphi = -\frac{1}{2}\Delta\varphi + V\varphi$ has nice diagonalization properties, namely a real, discrete, and bounded from below spectrum. Thus the resolution of the Schrödinger can be reduced to the computation of the eigenelements of H. Indeed, if the initial condition φ_0 of equation (6.1) can be decomposed as

$$\varphi_0 = \sum_{k=0}^{\infty} \lambda_k \psi_k,$$

where the $(\psi_k)_{k\geq 0}$ are the eigenvectors of H, then the solution φ_t of (6.1) can be written

$$\varphi_t = \sum_{k=0}^{\infty} \lambda_k e^{-iE_k t}\psi_k,$$

where the $(E_k)_{k\geq 0}$ are the respective eigenvalues of the $(\psi_k)_{k\geq 0}$. Practitioners are especially interested in the eigenvector associated to the smaller eigenvalue, called the *ground state*.

Depending on the type of particles considered, the natural space in which the wave function lies can be a strict subspace of $\mathbb{L}^2(\Omega)$. Indeed, if the considered system is composed of N indistinguishable bosonic particles, we only consider the space

$$\mathbb{L}^2_s\left(\left(\mathbb{R}^3\right)^N\right) = \left\{\varphi \in \mathbb{L}^2((\mathbb{R}^3)^N),\ \forall \sigma \in \mathfrak{S}_N,\ \varphi(x_1, \ldots, x_N) = \varphi(x_{\sigma(1)}, \ldots, x_{\sigma(N)})\right\}$$

of symmetric square integrable functions, and if the system is composed of N indistinguishable fermionic particles, we consider the space

$$\mathbb{L}^2_a\left(\left(\mathbb{R}^3\right)^N\right) = \left\{\varphi \in \mathbb{L}^2((\mathbb{R}^3)^N),\ \forall \sigma \in \mathfrak{S}_N,\ \varphi(x_1, \ldots, x_N) = \varepsilon(\sigma)\varphi(x_{\sigma(1)}, \ldots, x_{\sigma(N)})\right\}$$

of antisymmetric square integrable functions. This last example is fundamental in the domain of electronic structure computations, since electrons are fermions. The aim of electronic structure computation is to describe the quantum state of an entire molecule, that is to say, to compute the wave function associated to the physical system composed of every electrons and nuclei composing the molecule. However, an exact computation of this wave function is not affordable in practice, because the dimensionality of the corresponding problem is very high : for example for the molecule Li_8, the dimension is $8 \times (3+1) \times 3 = 96$, as each of the eight atoms of lithium is composed of three electrons and one nucleus evolving in three dimensions.

A natural way to simplify the problem is to fix the positions of the nuclei, and to consider only the electrons. This is known as the *Born-Oppenheimer approximation*. Since the electrons are indistinguishable fermionic particles, the wave function of the system lies in the space $\bigwedge_{i=1}^{N} \mathbb{L}^2(\mathbb{R}^3)$, where N, is the number of electrons considered. The potential undergone by the electrons depends on the location of the nuclei. It is typically of the form

$$V(x) = \sum_{i=1}^{N} V_1(x_i) + \sum_{1 \leq i < j \leq N} V_2(x_i - x_j),$$

where the first term in the right hand side describes the electron-electron interaction, while the second term describes the electron-nuclei interaction. Usually, $V_2(y)$ only depends on $|y|$, and V_1 is of the form

$$V_1(x) = -z_k \sum_{k=1}^{K} \rho_k \star \frac{1}{|x|},$$

where each $(\rho_k)_{k=1,\dots,K}$ describes the shape of a nucleus with charge z_k. The $(\rho_k)_{k=1,\dots,K}$ are either of the form δ_{y_k} (point-like nuclei), or are smooth functions with compact support and unit integral (smeared nuclei).

To compute the actual structure of the molecule, one has to compute the electronic structure when the nuclei are fixed in the correct position. One could find this correct configuration by computing the force exerted by the electrons on the nuclei. This force can be expressed as the derivative of the ground state energy with respect to the positions of the nuclei.

Our setting will be the following : we consider the Schrödinger operator H, defined on a dense subspace $\mathcal{D}(H)$ of some Hilbert space $\mathcal{H}$, the latter being a subspace of the space $\mathbb{L}^2(\Omega)$ of square-integrable functions, the choice of $\mathcal{H}$ depending on the considered problem. We assume that V is such that H is bounded from below, and that the smallest eigenvalue of H

$$E = \inf \left\{ \langle \varphi, H\varphi \rangle, \varphi \in \mathcal{D}(q_H), \langle \varphi, \varphi \rangle^2 = 1 \right\}, \tag{6.2}$$

is nondegenerate and isolated. Examples of such potential functions V can be found in [21]. In equation (6.2), $\mathcal{D}(q_H)$ denotes the domain of the following quadratic form, called *energy functional*,

$$\langle \varphi, H\varphi \rangle = \int_\Omega \frac{1}{2} |\nabla \varphi|^2 + V|\varphi|^2,$$

which satisfies the inclusions $\mathcal{D}(H) \subset \mathcal{D}(q_H) \subset \mathcal{H}$. In the expression of E, the infimum is reached for some normalized eigenvector ψ associated to E.

For the problem of computation of forces, we consider a family of self-adjoint operators $(H_\lambda)_{\lambda \in \mathbb{R}}$ depending smoothly on a real parameter. We assume that H_λ has a smallest eigenvalue E_λ, which is nondegenerate and isolated, and we denote by ψ_λ an associated normalized eigenvector chosen to depend smoothly on λ. One desires to compute the derivative $\partial_\lambda^0 E_\lambda$ of the fundamental energy with respect to λ, at $\lambda = 0$ (from now on, for a quantity Q_λ depending on λ, we denote by $\partial_\lambda^0 Q_\lambda$ the derivative of Q_λ with respect to λ at $\lambda = 0$).

This setting is also useful for the computation of the mean $\langle \psi, \mathbf{O}\psi \rangle = \int_\Omega \psi \mathbf{O} \psi$ of some observable $\mathbf{O}$ with respect to the ground state ψ of H. Indeed, if one considers the perturbed operators $H_\lambda = H + \lambda \mathbf{O}$, then the ground state ψ_0 of the operator H_0 taken at $\lambda = 0$ is equal to the ground state ψ of the operator H and one has, using the self-adjointness of H_λ :

$$\partial_\lambda^0 E_\lambda = \partial_\lambda^0 \langle \psi_\lambda, H_\lambda \psi_\lambda \rangle = 2 \langle \partial_\lambda^0 \psi_\lambda, E\psi_0 \rangle + \langle \psi_0, \mathbf{O}\psi_0 \rangle = \langle \psi_0, \mathbf{O}\psi_0 \rangle, \tag{6.3}$$

since the normalization condition on ψ_λ yields $\langle \partial_\lambda^0 \psi_\lambda, \psi_0 \rangle = 0$. Equation (6.3) is known as the *Hellmann-Feynman theorem* in the physics literature.

6.2 An asymptotic variance reduction method : the zero bias/zero variance principle

In practice, computing the smallest eigenvalue of H or H_λ by deterministic methods is made quite hard by the high-dimensionality of the considered problems : for example, for the Li_8 molecule in the Born-Oppenheimer approximation, the dimension of the problem is $8 \times 3 \times 3 = 72$. That is why probabilistic methods have been developed for treating this problem, especially Monte Carlo methods. The idea is to rewrite the desired quantity as the expectation of some random variable $f_\psi(X^\psi)$ where f_ψ is some function depending on the ground state ψ and X^ψ is a random variable whose distribution also depends on ψ. Then, one can perform the computation by sampling N independent realizations $(X^{n,\psi})_{n\in\{1,\dots,N\}}$ of the random variable X^ψ, where N is some large integer, and computing an empirical mean, namely :

$$\mathbb{E}\left[f_\psi(X^\psi)\right] = \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f_\psi(X^{n,\psi}).$$

Another possibility is to use a Markov process $(X_t^\psi)_{t\geq 0}$ which is ergodic with respect to the law of X^ψ. Then, one can perform the computation using the long time convergence of ergodic means :

$$\mathbb{E}\left[f_\psi(X^\psi)\right] = \lim_{T\to\infty} \frac{1}{T}\int_0^T f_\psi(X_t^\psi)\mathrm{d}t.$$

Of course, in most cases, the exact ground state cannot be known explicitly, so that one has to use an approximation f_φ of f_ψ using a *trial function* φ one thinks to be close to the exact ground state ψ, and sample a random variable X^φ whose distribution is an approximation of the distribution of X^ψ.

For those methods to be efficient, one needs to control two parameters measuring the error of the considered simulation, the bias

$$\mathbb{E}[f_\varphi(X^\varphi) - f_\psi(X^\psi)],$$

measuring the systematic error in the computation, and the variance

$$\mathbb{E}\left[\left(f_\varphi(X^\varphi) - E[f_\varphi(X^\varphi)]\right)^2\right],$$

measuring the statistical error. These quantities have to be estimated in terms of the error on guessing the ground state, namely $\delta\psi = \varphi - \psi$.

In this section, we present a method for computing the values E or $\partial_\lambda^0 E_\lambda$ with optimized estimators introduced in [7, 68] such that the bias and variance are of high order with respect to $\delta\psi$.

6.2.1 The variational Monte Carlo setting

The variational Monte Carlo (VMC) method relies on the fact that for a nice function φ, the overdamped Langevin dynamics

$$\mathrm{d}X_t = -\frac{\nabla\varphi}{\varphi}(X_t)\mathrm{d}t + \mathrm{d}W_t \tag{6.4}$$

is, under suitable conditions, ergodic with invariant distribution $\varphi^2/\langle\varphi,\varphi\rangle$. Thus, if X is a random variable with law $\varphi^2/\langle\varphi,\varphi\rangle$ it is possible to compute expectations of the form $\mathbb{E}[f(X)]$ using ergodic means $\frac{1}{T}\int_0^T f(X_t)\mathrm{d}t$ on the trajectory of $(X_t)_{t\geq 0}$.

Computation of the ground state energy

One can give a probabilistic interpretation to the ground state energy of a self adjoint operator H by using the probability density ψ^2. Indeed, one has

$$E = \int_\Omega \psi H \psi = \int_\Omega \frac{H\psi}{\psi}\psi^2 = \mathbb{E}\left[\frac{H\psi}{\psi}(X)\right],$$

where X is some random variable with distribution $\psi^2/\langle\varphi,\varphi\rangle$.

Using the variational Monte Carlo method with some trial function φ, one can compute the expectation $\mathbb{E}\left[\frac{H\varphi}{\varphi}(X)\right]$ where X is distributed according to the density $\varphi^2/\langle\varphi,\varphi\rangle$. Notice that this quantity is an overestimation of the actual fundamental energy E, since the latter is the smallest eigenvalue of H, so that

$$E = \langle\psi, H\psi\rangle \leq \frac{\langle\varphi, H\varphi\rangle}{\langle\varphi,\varphi\rangle} = \mathbb{E}\left[\frac{H\varphi}{\varphi}(X)\right].$$

An important property of this method is an asymptotic variance reduction property called the *zero bias/zero variance principle* in the physics literature, see for example [7, 68]. The point of the zero bias/zero variance principle is that, in the case where the trial probability $\varphi^2/\langle\varphi,\varphi\rangle$ is exactly the fundamental state ψ^2, the function $\frac{H\varphi}{\varphi}$ is constant and equal to the exact ground state energy E, so that an estimator of E based on the random variable $\frac{H\psi}{\psi}(X)$ has neither bias nor variance. When φ is an approximation of ψ, this is not true anymore, but one can estimate the bias and variance in terms of $\delta\psi = \frac{\varphi}{\sqrt{\langle\varphi,\varphi\rangle}} - \psi$.

Proposition 6.2.1 *Let X be a random variable distributed according to $\varphi^2/\langle\varphi,\varphi\rangle = (\psi+\delta\psi)^2$. Then the random variable $\frac{H\varphi}{\varphi}(X)$ is an estimator of the fundamental energy E whose bias and variance satisfy*

$$\left|\mathbb{E}\left[\frac{H\varphi}{\varphi}(X)\right] - E\right| + \mathbf{Var}\left(\frac{H\varphi}{\varphi}(X)\right) = \mathcal{O}(\delta\psi^2).$$

The $\mathcal{O}$ notation above being understood in the sense of the norm $\|\varphi\|_H^2 = \langle\varphi,\varphi\rangle + \langle H\varphi, H\varphi\rangle$.

Proof. The bias satisfies, using the self-adjointness of H and the relation $(H-E)\varphi = \sqrt{\langle\varphi,\varphi\rangle}\,(H-E)\,\delta\psi$,

$$\mathbb{E}\left[\frac{H\varphi}{\varphi}(X)\right] - E = \frac{\langle\varphi,(H-E)\varphi)\rangle}{\langle\varphi,\varphi\rangle} = \frac{\langle(H-E)\varphi,\delta\psi\rangle}{\sqrt{\langle\varphi,\varphi\rangle}} = \langle\delta\psi,(H-E)\delta\psi\rangle = \mathcal{O}(\delta\psi^2).$$

For the variance, notice that

$$\mathbb{E}\left[\left(\frac{H\varphi}{\varphi}(X) - \mathbb{E}\left[\frac{H\varphi}{\varphi}(X)\right]\right)^2\right] = \mathbb{E}\left[\left(\frac{H\varphi}{\varphi}(X) - E\right)^2\right] + \mathcal{O}(\delta\psi^2).$$

Moreover,

$$\begin{aligned}\mathbb{E}\left[\left(\frac{H\varphi}{\varphi}(X) - E\right)^2\right] = \frac{\langle H\varphi - E\varphi, H\varphi - E\varphi\rangle}{\langle\varphi,\varphi\rangle} &= \langle H\delta\psi - E\delta\psi, H\delta\psi - E\delta\psi\rangle \\ &= \mathcal{O}(\delta\psi^2),\end{aligned}$$

yielding the zero variance principle.

Computation of the energy derivative

A natural method to compute $\partial_\lambda^0 E_\lambda$ adapted from section 6.2.1 would be to use the equation

$$\begin{aligned}\partial_\lambda^0 E_\lambda = \partial_\lambda^0 \langle \psi_\lambda, H_\lambda \psi_\lambda \rangle &= \langle \psi_0, \partial_\lambda^0 H_\lambda \psi_0 \rangle + 2 \langle \partial_\lambda^0 \psi_\lambda, E \psi_0 \rangle \\ &= \mathbb{E}\left[\frac{\partial_\lambda^0 H_\lambda \psi_0}{\psi_0}(X) \right],\end{aligned}$$

where X is distributed according to ψ_0^2. Since the distribution ψ_0^2 is unknown, one could think of sampling a random variable X from a distribution $\varphi_0^2 / \langle \varphi_0, \varphi_0 \rangle$ close to ψ_0^2 and computing the mean of the random variable $\partial_\lambda^0 H_\lambda \varphi_0 / \varphi_0(X)$ through the variational Monte Carlo procedure. However to have a zero variance principle, the function $\partial_\lambda^0 H_\lambda \varphi_0 / \varphi_0$ has to be constant when the trial function φ_0 is exactly, up to a multiplicative constant, the ground state ψ_0, which is not the case in general. More precisely, when one samples from $(\psi_0 + \delta\psi_0)^2$, the bias is only of order $\mathcal{O}(\delta\psi_0)$ and the variance does not vanish as $\delta\psi_0$ goes to zero, unless the function $\partial_\lambda^0 H_\lambda \psi_0 / \psi_0$ is constant.

However, using a modified function $\mathcal{E}[\varphi_0, \partial_\lambda^0 \varphi_\lambda]$ instead of the naive choice $\partial_\lambda^0 H_\lambda \varphi_0 / \varphi_0$, one can recover the zero bias/zero variance principle. Indeed, let φ_λ be a trial function chosen so that there exists some constant c such that $c\varphi_\lambda$ is close to ψ_λ and such that $c\partial_\lambda^0 \varphi_\lambda$ is close to $\partial_\lambda^0 \psi_\lambda$. Performing the following computation

$$\begin{aligned}\partial_\lambda^0 \left(\frac{\langle \varphi_\lambda, H_\lambda \varphi_\lambda \rangle}{\langle \varphi_\lambda, \varphi_\lambda \rangle} \right) &= \partial_\lambda^0 \left(\int_\Omega \frac{H_\lambda \varphi_\lambda}{\varphi_\lambda} \frac{\varphi_\lambda^2}{\int_\Omega \varphi_\lambda^2} \right) \\ &= \int_\Omega \left(\frac{\partial_\lambda^0 H_\lambda \varphi_0}{\varphi_0} + \frac{H \partial_\lambda^0 \varphi_\lambda}{\varphi_0} + \frac{H\varphi_0}{\varphi_0} \frac{\partial_\lambda^0 \varphi_\lambda}{\varphi_0} \right) \frac{\varphi_0^2}{\int_\Omega \varphi_0^2} - 2 \frac{\int_\Omega \varphi_0 H \varphi_0 \int_\Omega \varphi_0 \partial_\lambda^0 \varphi_\lambda}{\left(\int_\Omega \varphi_0^2 \right)^2} \\ &= \mathbb{E}\left[\left(\frac{\partial_\lambda^0 H_\lambda \varphi_0}{\varphi_0} + \frac{H \partial_\lambda^0 \varphi_\lambda}{\varphi_0} + \frac{H\varphi_0}{\varphi_0} \frac{\partial_\lambda^0 \varphi_\lambda}{\varphi_0} - 2 \frac{\langle \varphi_0, H\varphi_0 \rangle}{\langle \varphi_0, \varphi_0 \rangle} \frac{\partial_\lambda^0 \varphi_\lambda}{\varphi_0} \right)(X) \right],\end{aligned}$$

one can see that the random variable

$$\mathcal{E}[\varphi_0, \partial_\lambda^0 \varphi_\lambda](X) = \left(\frac{\partial_\lambda^0 H_\lambda \varphi_0}{\varphi_0} + \frac{H \partial_\lambda^0 \varphi_\lambda}{\varphi_0} + \frac{H\varphi_0}{\varphi_0} \frac{\partial_\lambda^0 \varphi_\lambda}{\varphi_0} - 2 \frac{\langle \varphi_0, H\varphi_0 \rangle}{\langle \varphi_0, \varphi_0 \rangle} \frac{\partial_\lambda^0 \varphi_\lambda}{\varphi_0} \right)(X),$$

where X has distribution $\varphi_0^2 / \langle \varphi_0, \varphi_0 \rangle$, is an unbiased estimator of $\partial_\lambda^0 E_\lambda$ when $\varphi_\lambda = c\psi_\lambda$, for some constant c. Note that $\mathcal{E}[\varphi_0, \partial_\lambda^0 \varphi_\lambda]$ actually does not depend on the whole family (φ_λ), but only on φ_0 and its derivative $\partial_\lambda^0 \varphi_\lambda$.

Proposition 6.2.2 *Let* $\frac{\varphi_\lambda}{\sqrt{\langle \varphi_\lambda, \varphi_\lambda \rangle}} = \psi_\lambda + \delta\psi_\lambda$ *be some perturbation of* ψ_λ *depending smoothly on* λ *and let* X *be a random variable with distribution* $\varphi_0^2 / \langle \varphi_0, \varphi_0 \rangle$.

Then the random variable $\mathcal{E}[\varphi_0, \partial_\lambda^0 \varphi_\lambda](X)$ *is an estimator of* $\partial_\lambda^0 E_\lambda$ *whose variance and bias satisfy*

$$\left| \mathbb{E}\Big[\mathcal{E}[\varphi_0, \partial_\lambda^0 \varphi_\lambda](X) - \partial_\lambda^0 E_\lambda \Big] \right| = \mathcal{O}(\delta\psi_0^2 + \delta\psi_0 \partial_\lambda^0 \delta\psi_\lambda)$$

and

$$\mathbf{Var}\Big(\mathcal{E}[\varphi_0, \partial_\lambda^0 \varphi_\lambda](X) \Big) \leq K \left(1 + \left\| \frac{\partial_\lambda^0 \varphi_\lambda}{\varphi_0} \right\|_{\mathbb{L}^\infty(\Omega)} \right) \left\| \frac{\partial_\lambda^0 \varphi_\lambda}{\varphi_0} \right\|_{\mathbb{L}^\infty(\Omega)} \|\delta\psi_0\|_{H_\lambda}^2 + \mathcal{O}(\delta\psi_0^2 + \delta\psi_0 \partial_\lambda^0 \delta\psi_\lambda).$$

The $\mathcal{O}$ *above are considered in the norm* $\|\varphi\|_{H_\lambda}^2 = \langle \varphi, \varphi \rangle + \langle H\varphi, H\varphi \rangle + \langle \partial_\lambda^0 H_\lambda \varphi, \partial_\lambda^0 H_\lambda \varphi \rangle$.

Proof. Since ψ_λ and $\frac{\varphi_\lambda}{\sqrt{\langle \varphi_\lambda, \varphi_\lambda \rangle}} = \psi_\lambda + \delta\psi_\lambda$ are normalized, one has

$$\langle \psi_\lambda, \psi_\lambda \rangle = 1 = \langle \psi_\lambda + \delta\psi_\lambda, \psi_\lambda + \delta\psi_\lambda \rangle = \langle \psi_\lambda, \psi_\lambda \rangle + 2 \langle \delta\psi_\lambda, \psi_\lambda \rangle + \langle \delta\psi_\lambda, \delta\psi_\lambda \rangle,$$

yielding $\langle \delta\psi_\lambda, \psi_\lambda \rangle = -\frac{1}{2} \langle \delta\psi_\lambda, \delta\psi_\lambda \rangle$. As a consequence, it holds that

$$\begin{aligned}\mathbb{E}\Big[\mathcal{E}[\varphi_0,\partial_\lambda^0\varphi](X)\Big] = \partial_\lambda^0\left(\frac{\langle\varphi_\lambda,H_\lambda\varphi_\lambda\rangle}{\langle\varphi_\lambda,\varphi_\lambda\rangle}\right) =&\partial_\lambda^0\langle\psi_\lambda+\delta\psi_\lambda,H(\psi_\lambda+\delta\psi_\lambda)\rangle\\ =&\partial_\lambda^0 E_\lambda + 2\partial_\lambda^0\left(E_\lambda\langle\delta\psi_\lambda,\psi_\lambda\rangle\right)+\partial_\lambda^0\langle\delta\psi_\lambda,H_\lambda\delta\psi_\lambda\rangle\\ =&\partial_\lambda^0 E_\lambda - \partial_\lambda^0\left(E_\lambda\langle\delta\psi_\lambda,\delta\psi_\lambda\rangle\right)+\partial_\lambda^0\langle\delta\psi_\lambda,H_\lambda\delta\psi_\lambda\rangle\\ =&\partial_\lambda^0 E_\lambda + \mathcal{O}(\delta\psi_0^2+\delta\psi_0\partial_\lambda^0\delta\psi_\lambda),\end{aligned}$$

so that we have the zero bias principle.

For the zero variance principle, notice that the variance

$$\mathbb{E}\left[\left(\mathcal{E}[\varphi_0,\partial_\lambda^0\varphi_\lambda](X)-\partial_\lambda^0\left(\frac{\langle\varphi_\lambda,H_\lambda\varphi_\lambda\rangle}{\langle\varphi_\lambda,\varphi_\lambda\rangle}\right)\right)^2\right]$$

is $\langle\varphi_0,\varphi_0\rangle$ times the integral of the square of the function

$$\begin{aligned}\left(\mathcal{E}[\varphi_0,\partial_\lambda^0\varphi_\lambda](X)-\partial_\lambda^0\left(\frac{\langle\varphi_\lambda,H_\lambda\varphi_\lambda\rangle}{\langle\varphi_\lambda,\varphi_\lambda\rangle}\right)\right)\varphi_0 =&\partial_\lambda^0 H_\lambda\varphi_0 + H\partial_\lambda^0\varphi_\lambda - \frac{H\varphi_0}{\varphi_0}\partial_\lambda^0\varphi_\lambda - \varphi_0\partial_\lambda^0\left(\frac{\langle\varphi_\lambda,H_\lambda\varphi_\lambda\rangle}{\langle\varphi_\lambda,\varphi_\lambda\rangle}\right)\\ &+2\left(\frac{H\varphi_0}{\varphi_0}\partial_\lambda^0\varphi_\lambda-\frac{\langle\varphi_0,H\varphi_0\rangle}{\langle\varphi_0,\varphi_0\rangle}\partial_\lambda^0\varphi_\lambda\right) \qquad (6.5)\\ =&\partial_\lambda^0\left(\frac{H_\lambda\varphi_\lambda}{\varphi_\lambda}-\frac{\langle\varphi_\lambda,H_\lambda\varphi_\lambda\rangle}{\langle\varphi_\lambda,\varphi_\lambda\rangle}\right)\varphi_0+2\left(\frac{H\varphi_0}{\varphi_0}-\frac{\langle\varphi_0,H\varphi_0\rangle}{\langle\varphi_0,\varphi_0\rangle}\right)\partial_\lambda^0\varphi_\lambda.\end{aligned}$$

But one has

$$\frac{H_\lambda\varphi_\lambda}{\varphi_\lambda}-\frac{\langle\varphi_\lambda,H_\lambda\varphi_\lambda\rangle}{\langle\varphi_\lambda,\varphi_\lambda\rangle}=\frac{(H_\lambda-E_\lambda)\delta\psi_\lambda}{\varphi_\lambda}+\langle\delta\psi_\lambda,(E_\lambda-H_\lambda)\delta\psi_\lambda\rangle.$$

As a consequence, equation (6.5) rewrites

$$\begin{aligned}\left(\mathcal{E}[\varphi_0,\partial_\lambda^0\varphi_\lambda](X)-\partial_\lambda^0\left(\frac{\langle\varphi_\lambda,H_\lambda\varphi_\lambda\rangle}{\langle\varphi_\lambda,\varphi_\lambda\rangle}\right)\right)\varphi_0 =&\partial_\lambda^0\left(\frac{(H_\lambda-E_\lambda)\delta\psi_\lambda}{\varphi_\lambda}\right)\varphi_0+2\left(\frac{(H-E)\delta\psi_0}{\varphi_0}\right)\partial_\lambda^0\varphi_\lambda\\ &+\mathcal{O}(\delta\psi_0^2+\delta\psi_0\partial_\lambda^0\delta\psi_\lambda)\\ =&(H-E)\delta\psi_0\frac{\partial_\lambda^0\varphi_\lambda}{\varphi_0}+\partial_\lambda^0\left((H_\lambda-E_\lambda)\,\delta\psi_\lambda\right)\\ &+\mathcal{O}(\delta\psi+\partial_\lambda^0\delta\psi_\lambda) \qquad (6.6)\end{aligned}$$

Squaring equation (6.6), one can see that the variance of $\mathcal{E}[\varphi_0,\partial_\lambda^0\varphi_\lambda](X)$ is given by

$$\begin{aligned}\mathbf{Var}(\mathcal{E}[\varphi_0,\partial_\lambda^0\varphi_\lambda](X)) =&\int_\Omega\left((H-E)\delta\psi_0\frac{\partial_\lambda^0\varphi_\lambda}{\varphi_0}\right)\left((H-E)\delta\psi_0\frac{\partial_\lambda^0\varphi_\lambda}{\varphi_0}+2\partial_\lambda^0((H_\lambda-E_\lambda)\,\delta\psi_\lambda)\right)\\ &+\mathcal{O}(\delta\psi_0^2+\delta\psi_0\partial_\lambda^0\delta\psi_\lambda)\\ \leq&K\left(1+\left\|\frac{\partial_\lambda^0\varphi_\lambda}{\varphi_0}\right\|_{\mathbb{L}^\infty(\Omega)}\right)\left\|\frac{\partial_\lambda^0\varphi_\lambda}{\varphi_0}\right\|_{\mathbb{L}^\infty(\Omega)}\|\delta\psi_0\|_{H_\lambda}^2+\mathcal{O}(\delta\psi_0^2+\delta\psi_0\partial_\lambda^0\delta\psi_\lambda),\end{aligned}$$

yielding the desired zero variance principle.

6.2.2 The diffusion Monte Carlo setting

In practice, the variational Monte Carlo method presented in 6.2.1 is used to obtain a good trial function φ, but the actual computation of the ground state energy is made using the so-called diffusion Monte Carlo (DMC) method, which is unbiased, and thus more accurate.

This method relies on the fact that the solution to the partial differential equation

$$\partial_t \Phi_t(x) = -H\Phi_t(x) \text{ for } (t,x) \in [0,\infty) \times \Omega,$$

for a given initial condition $\Phi_0 = \varphi$, is equivalent to the function $e^{-Et}\psi \langle \psi, \varphi \rangle$ as t goes to infinity, provided $\langle \psi, \varphi \rangle \neq 0$. As a consequence, one can compute the ground state of H through the formula

$$E = \frac{\langle \psi, H\varphi \rangle}{\langle \psi, \varphi \rangle} = \lim_{t\to\infty} \frac{\langle \Phi_t, H\varphi \rangle}{\langle \Phi_t, \varphi \rangle}.$$

Indeed, the quantity $E(t) = \frac{\langle \Phi_t, H\varphi \rangle}{\langle \Phi_t, \varphi \rangle}$ is an overestimation of E and the convergence holds at exponential rate, since we assumed that E is a nondegenerate isolated eigenvalue of H.

From the Feynman-Kac formula, one has the following representation for Φ_t

$$\Phi_t(x) = \mathbb{E}\left[\varphi(x + W_t) e^{-\int_0^t V(x+W_s)\mathrm{d}s}\right], \tag{6.7}$$

where $(W_t)_{t\geq 0}$ is a standard Brownian motion, yielding a natural probabilistic method to compute Φ_t. However, in practice, the exponential factor yields a large variance, making this interpretation unsuitable for precise computations. Practitioners thus use some importance sampling method by considering the function $\tilde{f}_t(x) = \varphi(x)\Phi_t(x)/\langle \varphi, \varphi \rangle$ which gives another expression for the quantity $E(t)$:

$$E(t) = \frac{\int_\Omega \frac{H\varphi}{\varphi}(x)\tilde{f}_t(x)\mathrm{d}x}{\int_\Omega \tilde{f}_t(x)\mathrm{d}x}.$$

The function $\tilde{f}$ satisfies the following partial differential equation

$$\begin{cases} \partial_t \tilde{f} & = \frac{1}{2}\Delta\tilde{f} - \nabla\cdot(b\tilde{f}) - \frac{H\varphi}{\varphi}\tilde{f} \\ f_0 & = \varphi^2/\langle \varphi, \varphi \rangle \end{cases}, \tag{6.8}$$

where the drift term b is defined by

$$b(x) = \frac{\nabla\varphi}{\varphi}(x).$$

This partial differential equation has an interpretation as the density of a weighted diffusion. Indeed, if one defines h as

$$\int_\Omega g(x)h_t(x)\mathrm{d}x = \mathbb{E}\left[g(X_t)e^{-\int_0^t \frac{H\varphi}{\varphi}(X_s)\mathrm{d}s}\right],$$

where the process $(X_t)_{t\geq 0}$ follows the dynamics

$$\begin{cases} \mathrm{d}X_t & = b(X_t)\mathrm{d}t + \mathrm{d}W_t \\ X_0 & \text{distributed according to } \varphi^2/\langle \varphi, \varphi \rangle \end{cases}, \tag{6.9}$$

then the function h is a solution to equation (6.8). In this probabilistic interpretation, the weight $H\varphi/\varphi$ is the so called *local energy* which is a constant when φ is chosen to be the exact ground state. Thus one can expect $H\varphi/\varphi$ to have smaller fluctuations than V if the trial function φ is well chosen. Moreover the drift b is such that the distribution φ^2 is invariant for this process. For those reasons, one expects this interpretation to have a smaller variance than the interpretation (6.7).

However, the function h is not necessarily equal to $\tilde{f}$. This phenomenon is due to the fact that the sample paths of the diffusion (6.9) stay inside the connected components of the set $\{x \in \Omega, \varphi(x) \neq 0\}$. In fact, the function h can be explicited through the formula

$$h_t(x) = \varphi(x)^2 \mathbb{E}\left[e^{-\int_0^t \frac{H\varphi}{\varphi}(X_s)\mathrm{d}s}\right],$$

showing that $h_t(x)/\varphi(x)$ vanishes on the set of zeros of φ, whereas this property has no reason to hold for the function $\tilde{f}/\varphi = \Phi_t$.

We have a probabilistic method for computing the quantity

$$E_{DMC}(t) = \frac{\int_\Omega \frac{H\varphi}{\varphi}(x) h_t(x) \mathrm{d}x}{\int_\Omega h_t(x) \mathrm{d}x} = \frac{\mathbb{E}\left[\frac{H\varphi}{\varphi}(X_t) e^{-\int_0^t \frac{H\varphi}{\varphi}(X_s)\mathrm{d}s}\right]}{\mathbb{E}\left[e^{-\int_0^t \frac{H\varphi}{\varphi}(X_s)\mathrm{d}s}\right]}$$

which should be an approximation of $E(t)$. Actually, under some regularity assumptions,

$$\lim_{t\to\infty} E_{DMC}(t) = E_{DMC} = \inf\{\langle \chi, H\chi\rangle, \chi \in \mathcal{D}(q_H), \langle \chi, \chi\rangle = 1, \varphi^{-1}(0) \subset \chi^{-1}(0)\}$$

see for example [21]. Moreover, E_{DMC} is an overestimation of the ground state energy E, and the equality holds if and only if the sets of zeros of φ and ψ are equal. The set of zeros of ψ is known in the physics literature as the set of *nodes* of the ground state. The subsequent error in the computation of E is called the *fixed node approximation*. The diffusion Monte Carlo energy E_{DMC} can also be expressed as

$$E_{DMC} = \min_{n=1}^{N} \inf\{\langle \chi, H\chi\rangle, \chi \in \mathcal{D}(q_H), \langle \chi, \chi\rangle = 1, \chi \text{ supported in } \Omega^n\} \tag{6.10}$$

where the $(\Omega^n)_{n=1,\dots,N}$ are the connected components of $\{x \in \Omega, \varphi(x) \neq 0\}$.

In conclusion, the diffusion Monte Carlo method allows to sample, in the long time limit, a random variable whose law is $\psi^{FN}\varphi/\langle \psi^{FN}, \varphi\rangle$, where φ is a trial function, and ψ^{FN} is the ground state of a modified operator H^{FN} depending on the choice of φ. In particular, the function $\psi^{FN}\varphi$ has a constant sign. The operator H^{FN} is obtained as the restriction of the operator H to the space of function whose support is contained in one of the Ω^n achieving the minimum in (6.10).

The energy case

Let φ be a trial function such that $\langle \varphi, \psi\rangle \neq 0$, and denote $\delta\psi = \varphi/\langle \varphi, \psi\rangle - \psi$. To compute the ground state energy by the diffusion Monte Carlo method, we use the equality

$$E = \frac{\langle \varphi, H\psi\rangle}{\langle \varphi, \psi\rangle},$$

which implies that the ground state energy E is the expectation of the random variable $\frac{H\varphi}{\varphi}(X)$, where the random variable X has distribution $\varphi\psi/\langle \varphi, \psi\rangle$. As a consequence,

Proposition 6.2.3 *If X is distributed according to $\varphi\psi/\langle \varphi, \psi\rangle$, the random variable $\frac{H\varphi}{\varphi}(X)$ is an estimator of E, which has no bias, and whose variance satisfies*

$$\mathbf{Var}\left(\frac{H\varphi}{\varphi}(X)\right) \leq K \langle \varphi, \psi\rangle \left\|\frac{\psi}{\varphi}\right\|_{\mathbb{L}^\infty(\Omega)} \|\delta\psi\|_H^2,$$

for some constant K, with $\|\varphi\|_H^2 = \langle \varphi, \varphi\rangle + \langle H\varphi, H\varphi\rangle$.

Proof. For the bias, just notice that $\mathbb{E}\left(\frac{H\varphi}{\varphi}(X)\right) = \frac{\langle H\varphi, \psi\rangle}{\langle \varphi, \psi\rangle} = E$.

For the variance, it holds that

$$\mathbf{Var}\left(\frac{H\varphi}{\varphi}(X)\right) = \mathbb{E}\left[\left(\frac{H\varphi}{\varphi}(X) - E\right)^2\right] = \frac{\left\langle H\varphi - E\varphi, H\varphi\frac{\psi}{\varphi} - E\psi\right\rangle}{\langle\varphi,\psi\rangle}$$
$$= \langle\varphi,\psi\rangle\left\langle H\delta\psi - E\delta\psi, (H\delta\psi - E\delta\psi)\frac{\psi}{\varphi}\right\rangle,$$

concluding the proof.

Notice that Proposition 6.2.3 shows that the only bias encountered in the diffusion Monte Carlo method is due to the fixed node approximation.

The energy derivative case : a theoretical zero bias/zero variance estimator

For the energy derivative, perform the following derivative :

$$\partial_\lambda^0 E_\lambda = \partial_\lambda^0\left(\int_\Omega \frac{H_\lambda\varphi_\lambda}{\varphi_\lambda}\frac{\varphi_\lambda\psi_\lambda}{\int_\Omega\varphi_\lambda\psi_\lambda}\right)$$
$$= \int_\Omega\left(\frac{\partial_\lambda^0 H_\lambda\varphi_0}{\varphi_0} + \frac{H\partial_\lambda^0\varphi_\lambda}{\varphi_0} + \frac{H\varphi_0}{\varphi_0}\frac{\partial_\lambda^0\psi_\lambda}{\psi_0} - E\left(\frac{\partial_\lambda^0\varphi_\lambda}{\varphi_0} + \frac{\partial_\lambda^0\psi_\lambda}{\psi_0}\right)\right)\frac{\varphi_0\psi_0}{\int_\Omega\varphi_0\psi_0}.$$

As a consequence, one can obtain the derivative of the energy as the mean of the random variable

$$\mathcal{F}[\varphi_0,\partial_\lambda^0\varphi_\lambda](X) = \left(\frac{\partial_\lambda^0 H_\lambda\varphi_0}{\varphi_0} + \frac{H\partial_\lambda^0\varphi_\lambda}{\varphi_0} + \frac{H\varphi_0}{\varphi_0}\frac{\partial_\lambda^0\psi_\lambda}{\psi_0} - E\frac{\partial_\lambda^0\psi_\lambda}{\psi_0} - E\frac{\partial_\lambda^0\varphi_\lambda}{\varphi_0}\right)(X),$$

where X is distributed according to the density $\varphi_0\psi_0/\langle\varphi_0,\psi_0\rangle$. The statement of the zero bias/zero variance principle for this estimator is the following :

Proposition 6.2.4 *The estimator $\mathcal{F}[\varphi_0,\partial_\lambda^0\varphi_\lambda](X)$ has no bias*

$$\mathbb{E}[\mathcal{F}[\varphi_0,\partial_\lambda^0\varphi_\lambda](X)] = \partial_\lambda^0 E_\lambda,$$

and satisfies the following zero variance principle :

$$\mathbf{Var}\left(\mathcal{F}[\varphi_0,\partial_\lambda^0\varphi_\lambda](X)\right) \le K\langle\varphi_0,\psi_0\rangle\left(1 + \left\|\frac{\partial_\lambda^0\psi_\lambda}{\psi_0}\right\|_{\mathbb{L}^\infty(\Omega)}\right)$$
$$\times\left(\left\|\frac{\psi_0}{\varphi_0}\right\|_{\mathbb{L}^\infty(\Omega)} + \left\|\frac{\partial_\lambda^0\psi_\lambda}{\varphi_0}\right\|_{\mathbb{L}^\infty(\Omega)}\right)\left(\|\delta\psi_0\|_{H_\lambda} + \|\partial_\lambda^0\delta\psi_\lambda\|_{H_\lambda}\right)^2,$$

in the norm $\|\varphi\|_{H_\lambda}^2 = \langle\varphi,\varphi\rangle + \langle H\varphi, H\varphi\rangle + \langle\partial_\lambda^0 H_\lambda\varphi, \partial_\lambda^0 H_\lambda\varphi\rangle$, for some constant K.

Proof. By definition, the expectation of $\mathcal{F}[\varphi_0,\partial_\lambda^0\varphi_\lambda](X)$ is

$$\mathbb{E}\left[\mathcal{F}[\varphi_0,\partial_\lambda^0\varphi_\lambda](X)\right] = \partial_\lambda^0\left(\frac{\langle\varphi_\lambda, H_\lambda\psi_\lambda\rangle}{\langle\varphi_\lambda,\psi_\lambda\rangle}\right) = \partial_\lambda^0 E_\lambda.$$

The variance is given by the integral over Ω of

$$\left(\mathcal{F}[\varphi_0,\partial_\lambda^0\varphi_\lambda](X) - \partial_\lambda^0 E_\lambda\right)^2\frac{\varphi_0\psi_0}{\langle\varphi_0,\psi_0\rangle}$$
$$= \langle\varphi_0,\psi_0\rangle^{-1}\left(\partial_\lambda^0 H_\lambda\varphi_0 + H\partial_\lambda^0\varphi_\lambda + H\varphi_0\frac{\partial_\lambda^0\psi_\lambda}{\psi_0} - E\left(\partial_\lambda^0\psi_\lambda\frac{\varphi_0}{\psi_0} + \partial_\lambda^0\varphi_\lambda\right) - \partial_\lambda^0 E_\lambda\varphi_0\right)$$
$$\times\left(\partial_\lambda^0 H_\lambda\varphi_0\frac{\psi_0}{\varphi_0} + H\partial_\lambda^0\varphi_\lambda\frac{\psi_0}{\varphi_0} + \frac{H\varphi_0}{\varphi_0}\partial_\lambda^0\psi_\lambda - E\left(\partial_\lambda^0\psi_\lambda + \partial_\lambda^0\varphi_\lambda\frac{\psi_0}{\varphi_0}\right) - \partial_\lambda^0 E_\lambda\psi_0\right).$$

In this expression, the first factor in the right hand side reads

$$\partial_\lambda^0(H_\lambda\varphi_\lambda) - \partial_\lambda^0(E_\lambda\varphi_\lambda) + (H\varphi_0 - E\varphi_0)\,\frac{\partial_\lambda^0\psi_\lambda}{\psi_0} = \partial_\lambda^0(H_\lambda\delta\psi_\lambda - E_\lambda\delta\psi_\lambda) + (H\delta\psi_0 - E\delta\psi_0)\,\frac{\partial_\lambda^0\psi_\lambda}{\psi_0},$$

while the second factor reads

$$\left(\partial_\lambda^0(H_\lambda\varphi_\lambda) - \partial_\lambda^0(E_\lambda\varphi_\lambda)\right)\frac{\psi_0}{\varphi_0} + (H\varphi_0 - E\varphi_0)\,\frac{\partial_\lambda^0\psi_\lambda}{\varphi_0} = \partial_\lambda^0(H_\lambda\delta\psi_\lambda - E_\lambda\delta\psi_\lambda)\frac{\psi_0}{\varphi_0} + (H\delta\psi_0 - E\delta\psi_0)\,\frac{\partial_\lambda^0\psi_\lambda}{\varphi_0},$$

yielding the given variance for the estimator $\mathcal{F}[\varphi_0, \partial_\lambda^0\varphi_\lambda](X)$.

In practice, how to perform the energy derivative case ?

The weakness of the estimator $\mathcal{F}[\varphi_0, \partial_\lambda^0\varphi_\lambda](X)$ presented in section 6.2.2 is that it cannot be computed in practice, since the ground state ψ_λ is unknown. In a first time, a natural way to avoid this difficulty is to use the approximation $\frac{\partial_\lambda^0\varphi_\lambda}{\varphi_0} \simeq \frac{\partial_\lambda^0\psi_\lambda}{\psi_0}$, yielding the following modified estimator :

$$\tilde{\mathcal{F}}[\varphi_0, \partial_\lambda^0\varphi_\lambda](X) = \left(\frac{\partial_\lambda^0 H_\lambda\varphi_0}{\varphi_0} + \frac{H\partial_\lambda^0\varphi_\lambda}{\varphi_0} + \frac{H\varphi_0}{\varphi_0}\frac{\partial_\lambda^0\varphi_\lambda}{\varphi_0} - 2E\frac{\partial_\lambda^0\varphi_\lambda}{\varphi_0}\right)(X).$$

Proposition 6.2.5 *The estimator $\tilde{\mathcal{F}}[\varphi_0, \partial_\lambda^0\varphi_\lambda](X)$ satisfies the following zero bias principle*

$$\left|\mathbb{E}\left[\tilde{\mathcal{F}}[\varphi_0, \partial_\lambda^0\varphi_\lambda](X)\right] - \partial_\lambda^0 E_\lambda\right| \leq \langle\varphi_0, \psi_0\rangle\left(\|\delta\psi_0\|_H\|\partial_\lambda^0\delta\psi_\lambda\|_H\left\|\frac{\psi_0}{\varphi_0}\right\|_{\mathbb{L}^\infty(\Omega)} + \|\delta\psi_0\|_H^2\left\|\frac{\partial_\lambda^0\psi_\lambda}{\varphi_0}\right\|_{\mathbb{L}^\infty(\Omega)}\right),$$

in the norm $\|\varphi\|_H^2 = \langle\varphi,\varphi\rangle + \langle H\varphi, H\varphi\rangle$, and the following zero variance principle

$$\begin{aligned}\mathbf{Var}\left(\tilde{\mathcal{F}}[\varphi_0, \partial_\lambda^0\varphi_\lambda](X)\right) \leq{}& \left(\left|\frac{\partial_\lambda^0\langle\varphi_\lambda,\psi_\lambda\rangle}{\langle\varphi_0,\psi_0\rangle}\right| + \left\|\frac{\partial_\lambda^0\varphi_\lambda}{\varphi_0}\right\|_{\mathbb{L}^\infty(\Omega)}\right)\left\|\frac{\psi_0}{\varphi_0}\right\|_{\mathbb{L}^\infty(\Omega)}\left(\|\delta\psi_0\|_{H_\lambda} + \|\partial_\lambda^0\delta\psi_\lambda\|_{H_\lambda}\right)^2 \\ &+ \left|\mathbb{E}\left[\tilde{\mathcal{F}}[\varphi_0, \partial_\lambda^0\varphi_\lambda](X)\right] - \partial_\lambda^0 E_\lambda\right|^2.\end{aligned}$$

Proof. Notice that the error between $\tilde{\mathcal{F}}$ and $\mathcal{F}$ is given by

$$\tilde{\mathcal{F}}[\varphi_0, \partial_\lambda^0\varphi_\lambda](X) - \mathcal{F}[\varphi_0, \partial_\lambda^0\varphi_\lambda](X) = \left(\frac{H\varphi_0}{\varphi_0}(X) - E\right)\left(\frac{\partial_\lambda^0\varphi_\lambda}{\varphi_0} - \frac{\partial_\lambda^0\psi_\lambda}{\psi_0}\right),$$

whose mean is given by

$$\begin{aligned}\int_\Omega \frac{H\varphi_0 - E\varphi_0}{\langle\varphi_0,\psi_0\rangle\,\varphi_0}\left(\partial_\lambda^0\varphi_\lambda\psi_0 - \partial_\lambda^0\psi_\lambda\varphi_0\right) &= \int_\Omega \frac{H\delta\psi_0 - E\delta\psi_0}{\varphi_0}\left(\langle\varphi_0,\psi_0\rangle\left(\partial_\lambda^0\delta\psi_\lambda\psi_0 - \partial_\lambda^0\psi_\lambda\delta\psi_0\right) + \partial_\lambda^0\langle\varphi_\lambda,\psi_\lambda\rangle\frac{\varphi_0\psi_0}{\langle\varphi_0,\psi_0\rangle}\right) \\ &= \int_\Omega \frac{H\delta\psi_0 - E\delta\psi_0}{\varphi_0}\left(\langle\varphi_0,\psi_0\rangle\left(\partial_\lambda^0\delta\psi_\lambda\psi_0 - \partial_\lambda^0\psi_\lambda\delta\psi_0\right)\right),\end{aligned}$$

the last equality being deduced from the self adjointness of H. Since the estimator $\mathcal{F}[\varphi_0, \partial_\lambda^0\varphi_\lambda](X)$ is unbiased, we deduce the zero bias principle for the estimator $\tilde{\mathcal{F}}[\varphi_0, \partial_\lambda^0\varphi_\lambda](X)$.

From the zero bias property, the variance of $\tilde{\mathcal{F}}$ satisfies

$$\mathbf{Var}(\tilde{\mathcal{F}}[\varphi_0, \partial_\lambda^0\varphi_\lambda]) \leq \mathbb{E}\left[\left(\tilde{\mathcal{F}}[\varphi_0, \partial_\lambda^0\varphi_\lambda](X) - \partial_\lambda^0 E_\lambda\right)^2\right] + \left|\mathbb{E}\left[\tilde{\mathcal{F}}[\varphi_0, \partial_\lambda^0\varphi_\lambda](X)\right] - \partial_\lambda^0 E_\lambda\right|^2$$

As a consequence, the variance is given, up to a term of order two in $\delta\psi_0$ and $\partial_\lambda^0\delta\psi_\lambda$, by the integral over Ω of the function

$$\left(\partial_\lambda^0 H_\lambda\varphi_0 + H\partial_\lambda^0\varphi_\lambda + H\varphi_0\frac{\partial_\lambda^0\varphi_\lambda}{\varphi_0} - 2E\partial_\lambda^0\varphi_\lambda - \partial_\lambda^0 E_\lambda\varphi_0\right)^2 \frac{\psi_0}{\langle\varphi_0,\psi_0\rangle\,\varphi_0}$$
$$= \left(\partial_\lambda^0((H_\lambda - E_\lambda)\varphi_\lambda) + (H-E)\varphi_0\frac{\partial_\lambda^0\varphi_\lambda}{\varphi_0}\right)^2 \frac{\psi_0}{\langle\varphi_0,\psi_0\rangle\,\varphi_0}$$
$$= \left(\partial_\lambda^0(\langle\varphi_\lambda,\psi_\lambda\rangle\,(H_\lambda - E_\lambda)\delta\psi_\lambda) + \langle\varphi_0,\psi_0\rangle\,(H-E)\delta\psi_0\frac{\partial_\lambda^0\varphi_\lambda}{\varphi_0}\right)^2 \frac{\psi_0}{\langle\varphi_0,\psi_0\rangle\,\varphi_0},$$

which is of order two in $\delta\psi_0$ and $\partial_\lambda^0\delta\psi_\lambda$, yielding the zero variance principle.

As Proposition 6.2.4 shows, the computable zero bias/zero variance estimator $\tilde{\mathcal{F}}$ of the energy derivative is not unbiased anymore, which makes it much less accurate than the estimator $\mathcal{F}$, which is not computable. A way to avoid this problem is to consider a $\lambda-$dependent diffusion. Indeed, if $H_\lambda = -\Delta + V_\lambda$ is the fixed-node operator associated to the trial function φ_λ, and ψ_λ is the associated ground state, then

$$E_\lambda = \frac{\langle\psi_\lambda, H_\lambda\varphi_\lambda\rangle}{\langle\psi_\lambda,\varphi_\lambda\rangle} = \lim_{t\to\infty}\frac{\mathbb{E}\left[\frac{H_\lambda\varphi_\lambda}{\varphi_\lambda}(X_t^\lambda)e^{-\int_0^t\frac{H_\lambda\varphi_\lambda}{\varphi_\lambda}(X_s^\lambda)\mathrm{d}s}\right]}{\mathbb{E}\left[e^{-\int_0^t\frac{H_\lambda\varphi_\lambda}{\varphi_\lambda}(X_s^\lambda)\mathrm{d}s}\right]}, \tag{6.11}$$

where the process $(X_t^\lambda)_{t\geq 0}$ satisfies the dynamics

$$\mathrm{d}X_t^\lambda = -\nabla V_\lambda(X_t^\lambda)\mathrm{d}t + \mathrm{d}W_t.$$

Derivating equality (6.11) with respect to λ, and assuming one can exchange the limit and the derivative, one finds

$$\begin{aligned}
\partial_\lambda^0 E_\lambda &= \partial_\lambda^0\left(\frac{\langle\psi_\lambda, H_\lambda\varphi_\lambda\rangle}{\langle\varphi_\lambda,\psi_\lambda\rangle}\right) \qquad (6.12)\\
&= \lim_{t\to\infty}\mathbb{E}\left[\left(\partial_\lambda^0\left(\frac{H_\lambda\varphi_\lambda}{\varphi_\lambda}\right)(X_t^0) + \nabla\left(\frac{H\varphi_0}{\varphi_0}\right)(X_t^0)\cdot T_t\right.\right.\\
&\qquad \left.\left. - \int_0^t \partial_\lambda^0\left(\frac{H_\lambda\varphi_\lambda}{\varphi_\lambda}\right)(X_s^0) + \nabla\left(\frac{H\varphi_0}{\varphi_0}\right)(X_s^0)\cdot T_s\mathrm{d}s\right)e^{-\int_0^t\frac{H\varphi_0}{\varphi_0}(X_s^0)\mathrm{d}s}\right]\\
&\qquad \times\mathbb{E}\left[e^{-\int_0^t\frac{H\varphi_0}{\varphi_0}(X_s^0)\mathrm{d}s}\right]^{-1}\\
&\quad - \mathbb{E}\left[\frac{H\varphi_0}{\varphi_0}(X_t^0)e^{-\int_0^t\frac{H\varphi_0}{\varphi_0}(X_s^0)\mathrm{d}s}\right]\\
&\qquad \times\mathbb{E}\left[-\left(\int_0^t\partial_\lambda^0\left(\frac{H_\lambda\varphi_\lambda}{\varphi_\lambda}\right)(X_s^0) + \nabla\left(\frac{H\varphi_0}{\varphi_0}\right)(X_s^0)\cdot T_s\mathrm{d}s\right)e^{-\int_0^t\frac{H\varphi_0}{\varphi_0}(X_s^0)\mathrm{d}s}\right]\\
&\qquad \times\mathbb{E}\left[e^{-\int_0^t\frac{H\varphi_0}{\varphi_0}(X_t^0)\mathrm{d}s}\right]^{-2}.
\end{aligned}$$

In equation (6.12), $T_t = \partial_\lambda^0 X_t^\lambda$ is the so-called tangent vector of the process $(X_t^\lambda)_{t\geq 0}$. Equation (6.12) could be a good way to compute the zero bias/zero variance approximation of $\partial_\lambda^0 E_\lambda$, provided one can give a control on the tangent vector T_t.

6.3 Sensitivity of a diffusion with respect to some parameter

In view of sections 6.2.1 and 6.2.2, it is natural to consider some family of diffusion processes with a common diffusion term and a drift depending on some real parameter λ : let $(V_\lambda)_{\lambda\in\mathbb{R}}$ be a

family of smooth functions depending smoothly on λ, and consider the solutions of the following family of Langevin dynamics on Ω :

$$\begin{cases} \mathrm{d}X_t^\lambda &= -\frac{1}{2}\nabla V_\lambda(X_t^\lambda)\mathrm{d}t + \mathrm{d}W_t \\ X_0^\lambda &= X_0 \end{cases}, \quad \lambda \in \mathbb{R}, \tag{6.13}$$

where ∇ denotes the gradient with respect to the space variable, and $(W_t)_{t\geq 0}$ is a standard d-dimensional Brownian motion. The initial condition X_0 is some random variable which does not depend on λ.

Since we are going to use the long time behavior of this process, we assume that the trajectories never explode :

Assumption i *The potential V_λ is such that pathwise existence and uniqueness holds for the process $(X_t^\lambda)_{t\geq 0}$, for any positive time t.*

A sufficient condition for Assumption i to be satisfied is the following drift condition

$$x \cdot \nabla V_\lambda(x) \geq a|x|^2 + b$$

for some real numbers a and b.

We assume $\int_\Omega e^{-V_\lambda(x)}\mathrm{d}x < \infty$ for all λ, and, up to modifying V_λ by an additive function not depending on x, we can also assume that all $\pi_\lambda = e^{-V_\lambda(x)}\mathrm{d}x$ are probability measures. It is well known that the measures π_λ are invariant probability measures under the dynamics (6.13). This condition yields the ergodicity of the trajectories :

Proposition 6.3.1 *For all $\lambda \in \mathbb{R}$, the potential V_λ is such that the dynamics* (6.13) *is ergodic, with invariant measure π_λ. Namely, for a function f in $\mathbb{L}^1(\pi_\lambda)$, it holds almost surely, for any initial condition,*

$$\frac{1}{T}\int_0^T f(X_t^\lambda)\mathrm{d}t \underset{T\to\infty}{\to} \int_\Omega f\mathrm{d}\pi_\lambda.$$

Moreover, one has convergence in law of the time marginals : for any bounded function f, it holds that

$$\mathbb{E}[f(X_t^\lambda)] \underset{t\to\infty}{\to} \int_\Omega f\mathrm{d}\pi_\lambda.$$

Proof. The process $(X_t^\lambda)_{t\geq 0}$ is a diffusion whose trajectories are defined for any positive time, so that the existence of an invariant probability measure is equivalent to ergodicity ; see for example [41] for a proof of this fact in a more general setting. As a consequence, the process $(X_t^\lambda)_{t\geq 0}$ satisfies a law of large numbers, see [52].

For some Markov semigroup, the convergence in law of the time marginals holds as soon as the kernel of the generator is reduced to the space of constant variables. This follows from the spectral decomposition of the semigroup, see for example [8]. In our case, this condition is automatically satisfied, as one obviously sees on the following integration by parts

$$\int_{\mathbb{R}} u(x)\left(\frac{1}{2}\Delta u(x) - \frac{1}{2}\nabla V_0(x)\cdot\nabla u(x)\right)e^{-V_0(x)} = -\frac{1}{2}\int_{\mathbb{R}} |\nabla u(x)|^2 e^{-V_0(x)}\mathrm{d}x.$$

Our aim is now to compute the derivative

$$\lim_{\lambda\to 0}\frac{\int_\Omega f\mathrm{d}\pi_\lambda - \int_\Omega f\mathrm{d}\pi_0}{\lambda} = \partial_\lambda^0\left(\int_\Omega f\mathrm{d}\pi_\lambda\right), \tag{6.14}$$

for f in some family $\mathcal{C}$ of smooth functions. We need an assumption to ensure that $\int_\Omega f\mathrm{d}\pi_\lambda$ is differentiable :

Assumption ii *The potential V_λ is smooth enough with respect to λ so that for all functions f in $\mathcal{C}$, the quantity $\int_\Omega f \mathrm{d}\pi_\lambda$ is smooth with respect to λ.*

For example, this assumption is satisfied if the family V_λ is smooth with respect to λ and is such that for any f in $\mathcal{C}$ there exists some integrable nonnegative function g satisfying almost everywhere

$$\forall \lambda \in \mathbb{R},\ |\partial_\lambda V_\lambda(x) f(x) e^{-V_\lambda(x)}| \leq g(x).$$

According to the ergodicity assumption, it is natural to approximate (6.14) by the quantities $\partial_\lambda^0 \left(\frac{1}{T} \int_0^T f(X_t^\lambda) \mathrm{d}t \right)$ or $\partial_\lambda^0 \left(\mathbb{E}[f(X_t^\lambda)] \right)$ in the long time limit. In Proposition 6.3.5, we give a few conditions under which the equalities

$$\partial_\lambda^0 \left(\lim_{T\to\infty} \frac{1}{T} \int_0^T f(X_t^\lambda) \mathrm{d}t \right) = \lim_{T\to\infty} \frac{1}{T} \int_0^T \partial_\lambda^0 (f(X_t^\lambda)) \mathrm{d}t$$

and

$$\partial_\lambda^0 \left(\lim_{t\to\infty} \mathbb{E}\left[f(X_t^\lambda) \right] \right) = \lim_{t\to\infty} \mathbb{E}\left[\partial_\lambda^0 (f(X_t^\lambda)) \right]$$

hold. The function f being supposed smooth, the computation of this derivative can be done by computing the tangent vector, defined as the derivative of the trajectory X_t^λ with respect to the parameter λ :

$$T_t = \partial_\lambda^0 X_t^\lambda.$$

Indeed, we then have $\partial_\lambda^0 \left(f(X_t^\lambda) \right) = T_t \cdot \nabla f(X_t^0)$.

From the regularity of V_λ with respect to λ, standard differential equation arguments yield the following proposition :

Proposition 6.3.2 *For any $t \geq 0$, the function $\lambda \mapsto X_t^\lambda$ is almost surely differentiable. As a consequence, the definition of the tangent vector makes sense. Differentiating* (6.13) *with respect to λ, we obtain the following ordinary differential equation whose coefficients depend on X_t^0 :*

$$\begin{cases} \partial_t T_t &= -\frac{1}{2} \partial_\lambda^0 \nabla V_\lambda(X_t^0) - \frac{1}{2} \nabla^2 V_0(X_t^0) T_t \\ T_0 &= 0 \end{cases}. \tag{6.15}$$

The tangent vector T_t can take large values, since the second term in the right-hand side of (6.15) will provide exponential growth for T_t, when X_t^0 is close to a local maximum of V_0, or when X_t^0 crosses a saddle point of V_0.

We have the following expression of T_t as an integral :

Proposition 6.3.3 *For s, t in $[0, \infty)$, we define the so-called resolvent $R(s,t)$ of equation* (6.15)*, solution of the following ordinary differential equation :*

$$\begin{cases} \partial_t R(s,t) &= -\frac{1}{2} \nabla^2 V_0(X_t^0) R(s,t) \\ R(s,s) &= Id \end{cases}.$$

The resolvent satisfies the semi-group property : for any r, s, t in $[0, \infty)$, $R(s,t)R(r,s) = R(r,t)$. One can recover the tangent vector from the resolvent through the following formula :

$$T_t = -\frac{1}{2} \int_0^t R(s,t) \partial_\lambda^0 \nabla V_\lambda(X_s^0) \mathrm{d}s. \tag{6.16}$$

Proof. In view of the differential equations satisfied by T_t and $R(s,t)$, one has, since $R(t,0) = R(0,t)^{-1}$,

$$\begin{aligned}\partial_t(R(t,0)T_t) =& \frac{1}{2}R(t,0)\nabla^2 V_0(X_t^0)R(0,t)R(t,0)T_t - \frac{1}{2}R(t,0)\partial_\lambda^0\nabla V_\lambda(X_t^0) - \frac{1}{2}R(t,0)\nabla^2 V_0(X_t^0)T_t \\ =& -\frac{1}{2}R(t,0)\partial_\lambda^0\nabla V_\lambda(X_t^0).\end{aligned}$$

Integrating over $[0,t]$, one obtains

$$R(t,0)T_t = -\frac{1}{2}\int_0^t R(s,0)\partial_\lambda^0\nabla V_\lambda(X_s^0)\mathrm{d}s.$$

The result follows by using the semi-group property.

We need some assumption on the convexity of the potential for $R(s,t)$ to be well-behaved.

Assumption iii *The potential V_0 satisfies*

$$\int_\Omega \min \mathrm{Spec}\left(\nabla^2 V_0(x)\right) e^{-V_0(x)}\mathrm{d}x > 0. \tag{6.17}$$

While Assumption iii is automatically satisfied in dimension 1 from a simple integration by parts, this is not the case in higher dimension. If one applies the integration by parts formula in higher dimension, one only obtains that

$$\int_\Omega \nabla^2 V_0(x) e^{-V_0(x)}\mathrm{d}x$$

is a positive definite matrix, so that the minimum of its spectrum is positive. A counterexample is given by a tensor potential $V_0(x) = W(x_1) + \ldots + W(x_d)$ with a well chosen function W. Indeed, in this case the left hand side of equation (6.17) rewrites

$$\int_{\mathbb{R}^d} \min_i(W''(x_i)) e^{-\sum_i W(x_i)}\mathrm{d}x_1 \ldots \mathrm{d}x_d = \mathbb{E}\left[\min_{i=1}^d W''(X_i)\right],$$

where X_i are i.i.d random variables with distribution $e^{-W(x)}\mathrm{d}x$. If W is chosen so that W'' is bounded and has a strictly negative minimum, then the sequence $\left(\min_{i=1}^d W''(X_i)\right)$ converges in probability as d goes to infinity to the constant random variable $\min W''$, which is negative. Then from the dominated convergence theorem, the quantity $\mathbb{E}\left[\min_{i=1}^d W''(X_i)\right]$ is negative when d is large enough.

Lemma 6.3.4 *Under Assumptions i and iii, the matrix $R(0,t)$ almost surely converges to 0 as t goes to infinity, at exponential rate.*

Proof. For any vector x, one has

$$\begin{aligned}\partial_t\|R(0,t)x\|^2 =& -(R(0,t)x)^T\nabla^2 V_0(X_t^0)(R(0,t)x) \\ \leq& -\min \mathrm{Spec}\left(\nabla^2 V_0(X_t^0)\right)\|R(0,t)x\|^2.\end{aligned}$$

As a consequence, from Proposition 6.3.1, $\|R(0,t)x\|^2$ is smaller than $Ke^{-\int_0^t \min \mathrm{Spec}\left(\nabla^2 V_0(X_s^0)\right)\mathrm{d}s}$. By ergodicity, $-\int_0^t \min \mathrm{Spec}\left(\nabla^2 V_0(X_s^0)\right)\mathrm{d}s$ is equivalent to $-t\int_\Omega \min \mathrm{Spec}\left(\nabla^2 V_0(x)\right)e^{-V_0(x)}\mathrm{d}x$ as t goes to infinity, and thus $\|R(0,t)x\|^2$ converges to 0, with any rate $e^{-\beta t}$, $0 < \beta < \int_\Omega \min \mathrm{Spec}\left(\nabla^2 V_0(x)\right)e^{-V_0(x)}\mathrm{d}x$.

Assumption iii is not necessary for Lemma 6.3.4 to hold. Indeed, if the matrices A_t commute, for example in the case of a tensor potential $V_0(x) = W(x_1) + \ldots + W(x_d)$, Lemma 6.3.4 is always true, even in the cases when V_0 does not satisfy Assumption iii. However, some assumption is

needed, as there exists some family of matrices (A_t) converging in the Cesàro sense to a negative-definite matrix, for which the solution of $\partial_t R_t = A_t R_t$ does not vanish as t goes to infinity. An example of this phenomenon is given by

$$A_t = \Omega_t \begin{pmatrix} 1 & 0 \\ 0 & -3 \end{pmatrix} \Omega_t^T, \text{ where } \Omega_t = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}.$$

Indeed, the family $(A_t)_{t\in\mathbb{R}}$ converges in the Cesàro sense to $-Id$ as t goes to infinity, but the associated matrix $(R_t)_{t\in\mathbb{R}}$ diverges. To show this last point, consider the matrix $M_t = \Omega_t^T R_t$. Since $\partial_t \Omega_t = \Omega_t \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, it holds

$$\partial_t M_t = \begin{pmatrix} 1 & -1 \\ 1 & -3 \end{pmatrix} M_t.$$

As a consequence, $R_t = \Omega_t \exp\left(t \begin{pmatrix} 1 & -1 \\ 1 & -3 \end{pmatrix}\right) \Omega_t^T$. The eigenvalues of the matrix $\begin{pmatrix} 1 & -1 \\ 1 & -3 \end{pmatrix}$ are $-1-\sqrt{3}$ and $-1+\sqrt{3}$, the latter being positive, so that R_t diverges as t goes to infinity.

To prove the long time convergence, one will also need the couple (X_t^0, T_t) to be ergodic. However, the dynamics followed by this dynamics is not strongly elliptic, since it is driven by a one dimensional Brownian motion. Consequently, one needs another assumption to ensure convergence of the Cesàro means for the couple (X_t^0, T_t). We consider the following Hörmander condition for the dynamics of (X_t^0, T_t) :

Assumption iv *Assume that the potential V_λ is such that for any (x,t) in $\Omega \times \mathbb{R}^d$, the matrix*

$$\left(\partial_{ij} \left(\partial_\lambda^0 V_\lambda(x) + \sum_{k=1}^d \partial_k V_0(x) \cdot t_k \right) \right)_{i,j}$$

is nonsingular.

This condition is for example realized if V_0 is a quadratic potential and that the Hessian matrix of $\partial_\lambda^0 V_\lambda$ is everywhere nonsingular.

We can now prove our main result.

Proposition 6.3.5 *Let Assumption i, ii, iii and iv hold, and assume that $\partial_\lambda^0 \nabla V_\lambda$ is bounded and that the measure e^{-V_0} satisfies a Poincaré inequality. Let X_t^λ be a solution of* (6.13)*. Then,*

$$\lim_{T\to\infty} \partial_\lambda^0 \left(\frac{1}{T} \int_0^T f(X_t^\lambda) \mathrm{d}t \right) = \partial_\lambda^0 \left(\lim_{T\to\infty} \frac{1}{T} \int_0^T f(X_t^\lambda) \mathrm{d}t \right) = \partial_\lambda^0 \left(\int_\Omega f \mathrm{d}\pi_\lambda \right)$$

and

$$\lim_{t\to\infty} \partial_\lambda^0 \left(\mathbb{E}\left[f(X_t^\lambda) \right] \right) = \partial_\lambda^0 \left(\lim_{t\to\infty} \mathbb{E}\left[f(X_t^\lambda) \right] \right) = \partial_\lambda^0 \left(\int_\Omega f \mathrm{d}\pi_\lambda \right).$$

To prove Proposition 6.3.5, one first needs to know the long time limit of the trajectory and its tangent vector :

Lemma 6.3.6 *Under Assumption i and iii, if $\partial_\lambda^0 \nabla V_\lambda$ is bounded, as t goes to infinity, the couple (X_t^0, T_t) converges in law to the couple*

$$\left(Y_0, -\frac{1}{2} \int_0^\infty \tilde{R}(0,t) \partial_\lambda^0 \nabla V_\lambda(Y_t) \mathrm{d}t \right),$$

where Y follows the Langevin dynamics (6.13) *at $\lambda = 0$, but starting at equilibrium, that is to say, with Y_0 distributed according to $e^{-V_0(x)}\mathrm{d}x$, and $\tilde{R}(0,t)$ is the reversed resolvent associated to Y :*

$$\begin{cases} \partial_t \tilde{R}(s,t) &= -\frac{1}{2}\tilde{R}(s,t)\nabla^2 V_0(Y_t) \\ \tilde{R}(s,s) &= Id \end{cases}.$$

Proof. The integral $\int_0^\infty \tilde{R}(0,t)\partial_\lambda^0 \nabla V_\lambda(Y_t)\mathrm{d}t$ is almost surely well defined, from Proposition 6.3.4 and from the boundedness of $\partial_\lambda^0 \nabla V_\lambda$. To prove Lemma 6.3.6, we are going to use a time reversal argument.

We construct a coupling of the trajectory $(X_t^0)_{t\geq 0}$ with another process $(Y_t^\tau)_{t\geq\tau}$ following the Langevin dynamics (6.13), but being at equilibrium. Denote by p_t the density of the distribution of X_t^0, and define $\rho_t = \frac{p_t \wedge e^{-V_0}}{p_t}$. Let U and Z_t be mutually independent random variables which are independent of X_0^0 and of the Brownian motion $(W_t)_{t\geq 0}$ driving (X_t^0), such that U is uniformly distributed over $[0,1]$, and Z_t is distributed according to $C(e^{-V_0(x)} - p_t(x))^+\mathrm{d}x$, C being a normalization constant. We define the position of the process $(Y_t^\tau)_{t\geq 0}$ at time τ by $Y_\tau^\tau = X_\tau^0 \mathbf{1}_{U\leq \rho_\tau(X_\tau)} + Z_\tau \mathbf{1}_{U>\rho_\tau(X_t)}$, which is distributed according to π_0. One has $\mathbb{P}(Y_\tau^\tau \neq X_\tau^0) = \frac{1}{2}\|p_\tau(x)\mathrm{d}x - \pi_0\|_{TV}$. For $t > \tau$, let Y_t^τ evolve according to the dynamics (6.13) with Brownian motion $(W_t)_{t\geq 0}$, which admits π_0 as an invariant measure; Y_t^τ is thus distributed according to π_0 for all $t > \tau$, and satisfies $\mathbb{P}(\forall t \geq \tau, Y_t^\tau = X_t^0) = 1 - \frac{1}{2}\|p_\tau(x)\mathrm{d}x - \pi_0\|_{TV}$.

From Proposition 6.3.3, one has

$$(X_t^0, T_t) = \left(X_t^0, -\frac{1}{2}\int_0^t R(s,t)\partial_\lambda^0 \nabla V_\lambda(X_s^0)\mathrm{d}s\right),$$

and from the time reversibility of the dynamics (6.13), we have the equality in distribution, for $0 \leq \tau \leq t$,

$$\left(Y_t^\tau, -\frac{1}{2}\int_\tau^t R^\tau(s,t)\partial_\lambda^0 \nabla V_\lambda(Y_s^\tau)\mathrm{d}s\right) \overset{\mathcal{D}}{=} \left(Y_0, -\frac{1}{2}\int_0^{t-\tau} \tilde{R}(0,s)\partial_\lambda^0 \nabla V_\lambda(Y_s)\mathrm{d}s\right),$$

where R^τ is the resolvent associated to Y^τ :

$$\begin{cases} \partial_t R^\tau(s,t) &= -\frac{1}{2}\nabla^2 V_0(Y_t^\tau)R^\tau(s,t) \\ R^\tau(s,s) &= Id \end{cases}, \quad \forall s,t \geq \tau.$$

As a consequence, for any bounded Lipschitz continuous function φ,

$$\begin{aligned} &\mathbb{E}\left[\varphi(X_t^0, T_t) - \varphi\left(Y_0, -\frac{1}{2}\int_0^{t-\tau} \tilde{R}(0,s)\partial_\lambda^0 \nabla V_\lambda(Y_s)\mathrm{d}s\right)\right] \\ =&\mathbb{E}\left[\varphi(X_t^0, T_t) - \varphi\left(Y_t^\tau, -\frac{1}{2}\int_\tau^t R^\tau(s,t)\partial_\lambda^0 \nabla V_\lambda(Y_s^\tau)\mathrm{d}s\right)\right] \\ =&\mathbb{E}\left[\varphi(X_t^0, T_t) - \varphi\left(Y_t^\tau, -\frac{1}{2}\int_0^\tau \tilde{R}(s,t)\partial_\lambda^0 \nabla V_\lambda(X_s^0)\mathrm{d}s - \frac{1}{2}\int_\tau^t R^\tau(s,t)\partial_\lambda^0 \nabla V_\lambda(Y_s^\tau)\mathrm{d}s\right)\right] \\ &+\mathbb{E}\left[\varphi\left(Y_t^\tau, -\frac{1}{2}\int_0^\tau \tilde{R}(s,t)\partial_\lambda^0 \nabla V_\lambda(X_s^0)\mathrm{d}s - \frac{1}{2}\int_\tau^t R^\tau(s,t)\partial_\lambda^0 \nabla V_\lambda(Y_s^\tau)\mathrm{d}s\right)\right. \\ &\left.-\varphi\left(Y_t^\tau, -\frac{1}{2}\int_\tau^t R^\tau(s,t)\partial_\lambda^0 \nabla V_\lambda(Y_s^\tau)\mathrm{d}s\right)\right]. \end{aligned} \tag{6.18}$$

The first term in the right-hand side of (6.18) is smaller than $2\|\varphi\|_\infty \mathbb{P}(X_\tau^0 \neq Y_\tau^\tau) = \|\varphi\|_{\mathbb{L}^\infty(\Omega)}\|\pi_0 - p_\tau(x)\mathrm{d}x\|_{TV}$, which converges to 0 as τ goes to infinity. Indeed, from Cauchy-Schwarz inequality

$$\|p_\tau(x)\mathrm{d}x - \pi_0\|_{TV} \leq \left(\int_\Omega \left(1 - \frac{p_\tau(x)}{e^{-V_0(x)}}\right)^2 e^{-V_0(x)}\mathrm{d}x\right)^{1/2} \left(\int_\Omega e^{-V_0(x)}\mathrm{d}x\right)^{1/2},$$

and from Poincaré inequality, the quantity $\int_\Omega (1 - \frac{p_\tau}{e^{-V_0}})^2 e^{-V_0}$ vanishes as τ goes to infinity.

The second term in (6.18) is smaller than $\mathbb{E}\left[2\|\varphi\|_{\mathbb{L}^\infty(\Omega)} \wedge \left(\frac{1}{2}|\varphi|_{\mathrm{Lip}} \int_0^\tau \|\tilde{R}(s,t)\partial_\lambda^0 \nabla V_\lambda(X_s^0)\|\mathrm{d}s\right)\right]$. A direct adaptation of Lemma 6.3.4 shows that $\tilde{R}(s,t)$ vanishes as t goes to infinity, yielding from Lebesgue's theorem that the second term in (6.18) vanishes as t goes to infinity. One can thus conclude the proof by letting t and then τ go to infinity.

Proof (Proof of Proposition 6.3.5). Denote by μ the distribution of the initial condition X_0^λ. From assumption iv, which is a hypoellipticity assumption on the process (X_t^0, T_t), the convergence in distribution in Lemma 6.3.6 actually holds almost surely in the Cesàro sense. Thus, one can write

$$\begin{aligned}
\lim_{T\to\infty} \partial_\lambda^0 \left(\frac{1}{T}\int_0^T f(X_t^\lambda)\mathrm{d}t\right) &= \lim_{T\to\infty} \frac{1}{T}\int_0^T \nabla f(X_t^0)\cdot T_t \mathrm{d}t \\
&= -\frac{1}{2}\mathbb{E}_{\pi_0}\left[\nabla f(X_0^0)\cdot \int_0^\infty \tilde{R}(0,t)\partial_\lambda^0 \nabla V_\lambda(X_t^0)\mathrm{d}t\right] \\
&= -\frac{1}{2}\int_\Omega \nabla f(x)\cdot \int_0^\infty \mathbb{E}_x\left[\tilde{R}(0,t)\partial_\lambda^0 \nabla V_\lambda(X_t^0)\right]\mathrm{d}t\ e^{-V_0(x)}\mathrm{d}x \\
&= \frac{1}{2}\int_\Omega f(x)\nabla\cdot\left(\int_0^\infty \mathbb{E}_x\left[\tilde{R}(0,t)\partial_\lambda^0 \nabla V_\lambda(X_t^0)\right]\mathrm{d}t\ e^{-V_0(x)}\right)\mathrm{d}x.
\end{aligned}$$

The function $\mathbb{E}_x\left[\tilde{R}(0,t)\partial_\lambda^0 \nabla V_\lambda(X_t^0)\right]$ has an interpretation in terms of partial differential equation. Indeed, let v be the solution to the following partial differential equation :

$$\begin{cases} \partial_t v &= \frac{1}{2}\Delta v - \frac{1}{2}\nabla v\cdot\nabla V_0 \\ v_0 &= \partial_\lambda^0 V_\lambda \end{cases}.$$

Then ∇v satisfies the following equation :

$$\begin{cases} \partial_t u &= \frac{1}{2}\Delta u - \frac{1}{2}\nabla u\cdot\nabla V_0 - \frac{1}{2}\nabla^2 V_0 u \\ u_0 &= \partial_\lambda^0 \nabla V_\lambda \end{cases}. \tag{6.19}$$

In the above formula, Δu denotes the vector whose coordinates are Δu^i, and $\nabla u\cdot\nabla V_0$ denotes the vector whose coordinates are $\nabla u^i\cdot\nabla V_0$. By the Feynman-Kac formula,

$$\nabla v_t(x) = \mathbb{E}_x\left[\tilde{R}(0,t)\partial_\lambda^0 \nabla V_\lambda(X_t^0)\right].$$

Indeed, let u satisfy equation (6.19) and apply Itō's formula to the process $\tilde{R}(0,s)u_{t-s}^i(X_s^0)$. One obtains that, for any i in $\{1,\dots,d\}$,

$$\begin{aligned}
&\tilde{R}(0,t)u_0^i(X_t^0) - u_t^i(X_0^0) \\
&-\int_0^t \tilde{R}(s,t)\left(-\frac{1}{2}\nabla^2 V_0 u_{t-s}^i - \partial_t u_{t-s}^i + \frac{1}{2}\Delta u_{t-s}^i - \frac{1}{2}\nabla V_0\cdot\nabla u_{t-s}^i\right)(X_s)\mathrm{d}s
\end{aligned} \tag{6.20}$$

is a martingale. Since u satisfies (6.19), then the integral in (6.20) is zero, and $\tilde{R}(0,t)u_0^i(X_t^0) - u_t^i(X_0^0)$ is a martingale. Since $u_0 = \partial_\lambda^0 \nabla V_0$ and $X_0^O = x$ almost surely under $\mathbb{E}_x$, one obtains

$$u_t(x) = \mathbb{E}_x\left[\tilde{R}(0,t)\partial_\lambda^0 \nabla V(X_t)\right].$$

As a consequence,

$$\begin{aligned}\lim_{T\to\infty}\partial_\lambda^0\left(\frac{1}{T}\int_0^T f(X_t^\lambda)\mathrm{d}t\right) &= \int_\Omega f(x)\nabla\cdot\left(\int_0^\infty \nabla v_t(x)\mathrm{d}t\; e^{-V_0(x)}\right)\mathrm{d}x\\ &= \int_\Omega f(x)\int_0^\infty e^{-V_0(x)}\left(\Delta v_t(x)-\nabla v_t(x)\cdot\nabla V_0(x)\right)\mathrm{d}t\;\mathrm{d}x\\ &= \int_\Omega f(x)\int_0^\infty \partial_t v_t(y)\mathrm{d}t\; e^{-V_0(x)}\mathrm{d}x.\end{aligned}$$

The function v_t converges as t goes to infinity to the mean value of v_0 :

$$\lim_{t\to\infty} v_t(x) = \int_\Omega \partial_\lambda^0 V_\lambda(y)e^{-V_0(y)}\mathrm{d}y = \partial_\lambda^0\left(\int_\Omega e^{-V_0(y)}\mathrm{d}y\right) = 0.$$

As a consequence,

$$\begin{aligned}\lim_{T\to\infty}\partial_\lambda^0\left(\frac{1}{T}\int_0^T f(X_t^\lambda)\mathrm{d}t\right) = -\int_\Omega f(x)\partial_\lambda^0 V_\lambda(x)e^{-V_0(x)}\mathrm{d}x &= \partial_\lambda^0\left(\int_\Omega f(x)e^{-V_\lambda(x)}\mathrm{d}x\right)\\ &= \partial_\lambda^0\left(\lim_{T\to\infty}\frac{1}{T}\int_0^T f(X_t^\lambda)\mathrm{d}t\right),\end{aligned}$$

completing the proof for ergodic means. The proof is essentially the same for the convergence of the time marginals, up to replacing convergence of ergodic means by convergence of the marginals.

6.4 Numerical computation and variance reduction through particle filtering

Numerically, the computation of (6.14) through the expression

$$\partial_\lambda^0\left(\int_\Omega f\mathrm{d}\pi_\lambda\right) = \mathbb{E}[T_t\cdot\nabla f(X_t^0)]$$

can be efficiently performed provided T_t has a small variance. This is ensured under some assumptions :

Proposition 6.4.1 *Assume that V_0 is α-convex, for some positive α, and that $\partial_\lambda^0\nabla V_\lambda$ is bounded. Then, T_t has a bounded variance uniformly in time.*

Proof. From Proposition 6.3.3, T_t has an explicit form :

$$T_t = -\frac{1}{2}\int_0^t R(s,t)\partial_\lambda^0\nabla V_\lambda(X_s^0)\mathrm{d}s.$$

Since V_0 is $\alpha-$convex $R(s,t)\leq e^{-\alpha(t-s)}$ holds. It is sufficient to control the expectation of T_t^2 in order to control $\mathbf{Var}(T_t)$. One has

$$\begin{aligned}\mathbb{E}[T_t^2] &= \frac{1}{4}\mathbb{E}\left[\left(\int_0^t R(s,t)\partial_\lambda^0\nabla V_\lambda(X_s^0)\mathrm{d}s\right)^2\right]\\ &\leq \frac{1}{4}\mathbb{E}\left[\left(K\int_0^t e^{-\alpha(t-s)}\mathrm{d}s\right)^2\right] < \infty,\end{aligned}$$

so that T_t has a finite variance.

In practice, the potential V_0 is not convex, and the second term in the right hand side of (6.15) will provide exponential growth for T_t in the neighborhood of the saddle points and local maxima of V_0. As a consequence, the vector T_t will have a large variance.

One can reduce the variance of T_t by using a particle filter, based on the following Feynman-Kac representation :

$$\mathbb{E}[\nabla f(X_t)\cdot T_t] = \mathbb{E}\left[\nabla f(X_{t_n})\frac{T_{t_n}}{\|T_{t_n}\|}\|T_{t_0}\|\prod_{k=1}^{n}\frac{\|T_{t_k}\|}{\|T_{t_{k-1}}\|}\right]$$

where $0 < t_0 \leq t_1 \leq \ldots \leq t_n = t$. For a detailed overview of Feynman-Kac formulæ, see [27].

A possible way to reduce the variance of the vector T_t is to use a resampling procedure. The principle is to simulate a large number N of copies of the system, and from time to time to kill some of the copies and to multiply other ones. The aim of this resampling is to have more precision on the copies having a large tangent vector by reproducing them. As a consequence at each resampling, N new copies of the system will be created and evolve independently, but starting form position picked at random amongst the positions before the resampling, the positions with the larger T_t being the most likely to be chosen. One can think of two ways of choosing at what time performing the resampling : either with a pre-determined schedule, or when there is too much discrepancy between the sizes of the different tangent vectors.

6.4.1 Particle filtering with deterministic time grid

In this section we present a resampling algorithm where the resamplings occur at each step of a pre-determined schedule $0 < t_0 < t_1 < \ldots$.

1. **Initialization :** Up to time t_0, run N independent copies of the stochastic differential equation (6.13) for $\lambda = 0$ starting at x, and of the associated tangent equation (6.15) starting at 0. Let $(\tilde{X}_0^i, \tilde{T}_0^i)_{i=0,\ldots,N}$ be the result of those simulations, let $T_0^{0,i} = \tilde{T}_0^i$, and set $k = 1$.
2. **Iteration :**
   a) Up to time t_k, run N independent copies of the stochastic differential equation (6.13) for $\lambda = 0$ starting at $\tilde{X}_{k-1}^i$, and of the associated tangent equation (6.15) starting at $\tilde{T}_{k-1}^i$. Let $(X_k^i, T_k^i)_{i=0,\ldots,N}$ be the result of those simulations.
   b) Randomly and independently choose $(\tilde{X}_k^i, \tilde{T}_k^i, \|T_k^{0,i}\|)_{i=1,\ldots,N}$ amongst the triples $(X_k^i, T_k^i, \|T_{k-1}^{0,i}\|)_{i=1,\ldots,N}$, each triple having a weight proportional to $\frac{\|T_k^i\|}{\|T_{k-1}^i\|}$.
   c) Increment k and start again from step (a).

The purpose of the number $\|T_k^{0,i}\|$ is to keep track of the norm at time t_0 of the particle labelled i. One cannot set $t_0 = 0$ since the initial condition of the tangent vector is $T_0 = 0$, preventing us from dividing by $\|T_{t_0}\|$. This algorithm is designed to be unbiased, up to a computable renormalization :

Proposition 6.4.2 *For all $k \geq 0$, it holds that*

$$\mathbb{E}\left[\prod_{q=1}^{k}\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\|T_q^i\|}{\|\tilde{T}_{q-1}^i\|}\right)\frac{1}{N}\sum_{i=1}^{N}\nabla f(\tilde{X}_k^i)\cdot\frac{\tilde{T}_k^i}{\|\tilde{T}_k^i\|}\|T_k^{0,i}\|\right] = \mathbb{E}[\nabla f(X_{t_k})\cdot T_{t_k}]. \tag{6.21}$$

Proof. We are actually going to prove the following result : for any j in $\{0,\ldots,k\}$

$$\begin{aligned}&\mathbb{E}\left[\prod_{q=1}^{k}\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\|T_q^i\|}{\|\tilde{T}_{q-1}^i\|}\right)\frac{1}{N}\sum_{i=1}^{N}\nabla f(\tilde{X}_k^i)\cdot\frac{\tilde{T}_k^i}{\|\tilde{T}_k^i\|}\|T_k^{0,i}\|\right] \\ =&\mathbb{E}\left[\prod_{q=1}^{j}\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\|T_q^i\|}{\|\tilde{T}_{q-1}^i\|}\right)\frac{1}{N}\sum_{i=1}^{N}P_{t_k-t_j}\varphi\left(\tilde{X}_j^i,\tilde{T}_j^i\right)\frac{\|T_j^{0,i}\|}{\|\tilde{T}_j^i\|}\right],\end{aligned} \tag{6.22}$$

where P_t denotes the semi-group of the Markov Process $(X_t, T_t)_{t\geq 0}$, and φ denotes the function $\varphi(x,t) = \nabla f(x)\cdot t$. Equation (6.22) gives the desired result for $j=0$.

Equation (6.22) is trivial for $j=k$. Now, let $0<j\leq k$. One has, because of the choice of weights for the resampling,

$$\begin{aligned}
&\mathbb{E}\left[\prod_{q=1}^{j}\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\|T_q^i\|}{\|\tilde{T}_{q-1}^i\|}\right)\frac{1}{N}\sum_{i=1}^{N}P_{t_k-t_j}\varphi(\tilde{X}_j^i,\tilde{T}_j^i)\frac{\|T_j^{0,i}\|}{\|\tilde{T}_j^i\|}\right]\\
=&\mathbb{E}\left[\prod_{q=1}^{j-1}\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\|T_q^i\|}{\|\tilde{T}_{q-1}^i\|}\right)\frac{1}{N}\sum_{i=1}^{N}P_{t_k-t_j}\varphi(X_j^i,T_j^i)\frac{\|T_{j-1}^{0,i}\|}{\|T_j^i\|}\times\frac{\|T_j^i\|}{\|\tilde{T}_{j-1}^i\|}\right]\\
=&\mathbb{E}\left[\prod_{q=1}^{j-1}\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\|T_q^i\|}{\|\tilde{T}_{q-1}^i\|}\right)\frac{1}{N}\sum_{i=1}^{N}P_{t_k-t_j}\varphi(X_j^i,T_j^i)\frac{\|T_{j-1}^{0,i}\|}{\|\tilde{T}_{j-1}^i\|}\right].
\end{aligned}$$

Now, applying the Markov property, it holds that

$$\begin{aligned}
&\mathbb{E}\left[\prod_{q=1}^{j-1}\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\|T_q^i\|}{\|\tilde{T}_{q-1}^i\|}\right)\frac{1}{N}\sum_{i=1}^{N}P_{t_k-t_j}\varphi(X_j^i,T_j^i)\frac{\|T_{j-1}^{0,i}\|}{\|\tilde{T}_{j-1}^i\|}\right]\\
=&\mathbb{E}\left[\prod_{q=1}^{j-1}\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\|T_q^i\|}{\|\tilde{T}_{q-1}^i\|}\right)\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[P_{t_k-t_j}\varphi(X_j^i,T_j^i)|(X_q^i,T_q^i)_{0\leq q\leq j-1}\right]\frac{\|T_{j-1}^{0,i}\|}{\|\tilde{T}_{j-1}^i\|}\right]\\
=&\mathbb{E}\left[\prod_{q=1}^{j-1}\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\|T_q^i\|}{\|\tilde{T}_{q-1}^i\|}\right)\frac{1}{N}\sum_{i=1}^{N}P_{t_j-t_{j-1}}P_{t_k-t_j}\varphi(X_{j-1}^i,T_{j-1}^i)\frac{\|T_{j-1}^{0,i}\|}{\|\tilde{T}_{j-1}^i\|}\right]\\
=&\mathbb{E}\left[\prod_{q=1}^{j-1}\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\|T_q^i\|}{\|\tilde{T}_{q-1}^i\|}\right)\frac{1}{N}\sum_{i=1}^{N}P_{t_k-t_{j-1}}\varphi(X_{j-1}^i,T_{j-1}^i)\frac{\|T_{j-1}^{0,i}\|}{\|\tilde{T}_{j-1}^i\|}\right].
\end{aligned}$$

As a consequence equation (6.22) holds by induction.

6.4.2 Particle filtering with random time grid

In section 6.4.1, the times at which a resampling was performed were chosen *a priori*. It is possible to adapt our algorithm in order to resample at better chosen times, that may depend on the result on the trajectory. The purpose of the resampling is to reduce variance, when a particle has a dominant weight. It could thus be a good idea to resample when the weights have become too unbalanced.

Let $(w_i)_{i=1\dots N}$ be a family of weights (that is $w_i\geq 0$ and $\sum_i w_i = 1$). Then the entropy of the weights

$$H(w) = \sum_i w_i \log(N w_i)$$

is a nonnegative number, thanks to Jensen's inequality, and is equal to zero if and only if $w_i = \frac{1}{N}$ for all i. $H(w)$ is maximal when all w_i are 0 except for one index i_0 such that $w_{i_0}=1$, and then $H(w) = \log N$. Hence the entropy is a good measure for the "nonuniformity" of a distribution.

Consequently, we can perform the algorithm of section 6.4.1, but with random times $(t_i)_{i\geq 1}$ corresponding to instants when the relative entropy of the weights is too large, for example, as soon as the entropy is greater than $\frac{1}{2}\log N$. More precisely, t_0 is a deterministic positive time, and

$$t_{k+1} = \inf\left\{t\geq t_k, H\left(\left(\frac{\|T_k^i\|}{\|\tilde{T}_{k-1}^i\|}\right)_{i=1\dots,N}\right)\geq\frac{1}{2}\log N\right\}.$$

6.4.3 Particle merging

The couple (T_t, X_t) is measurable with respect to the $\sigma-$field generated by the random variables X_s, and $\mathcal{W}_s^t := (W_r)_{s\leq r\leq t}$. As a consequence, one can write X_t and T_t as measurable functions of $(X_s, \mathcal{W}_s^t)$. We denote

$$X_t = F(X_s, \mathcal{W}_s^t), \text{ and } T_t = G(X_s, \mathcal{W}_s^t)T_s + H(X_s, \mathcal{W}_s^t.)$$

Indeed, in view of the equality (6.16) and using the semigroup property of $R(s,t)$, one can see that T_t is an affine function of T_s. Conditioning with respect to $(X_s, \mathcal{W}_s^t)$, one has

$$\begin{aligned}\mathbb{E}[\nabla f(X_t)\cdot T_t] =& \mathbb{E}\left[\nabla f(F(X_s,\mathcal{W}_s^t))\cdot\left(G(X_s,\mathcal{W}_s^t)T_s + H(X_s,\mathcal{W}_s^t)\right)\right]\\ =& \mathbb{E}\left[\mathbb{E}\left[\nabla f(F(X_s,\mathcal{W}_s^t))\cdot\left(G(X_s,\mathcal{W}_s^t)T_s + H(X_s,\mathcal{W}_s^t)\right)|X_s,\mathcal{W}_s^t\right]\right]\\ =& \mathbb{E}[\nabla f(F(X_s,\mathcal{W}_s^t))\cdot\left(G(X_s,\mathcal{W}_s^t)\mathbb{E}[T_s|X_s,\mathcal{W}_s^t] + H(X_s,\mathcal{W}_s^t)\right)].\end{aligned}$$

From the independence of T_s and $\mathcal{W}_s^t$, one can finally write

$$\mathbb{E}[\nabla f(X_t)\cdot T_t] = \mathbb{E}[\nabla f(F(X_s,\mathcal{W}_s^t))\cdot\left(G(X_s,\mathcal{W}_s^t)\mathbb{E}[T_s|X_s] + H(X_s,\mathcal{W}_s^t)\right)].$$

As a consequence, if one defines $\tilde{T}_t$ to be the solution to

$$\begin{cases}\partial_t\tilde{T}_t &= -\frac{1}{2}\partial_\lambda^0\nabla V_\lambda(X_t^0) - \frac{1}{2}\nabla^2 V_0(X_t^0)\tilde{T}_t\\ \tilde{T}_s &= \mathbb{E}[T_s|X_s]\end{cases}, \ t\geq s$$

then, for any $t \geq s$,

$$\mathbb{E}[\nabla f(X_t)\cdot T_t] = \mathbb{E}[\nabla f(X_t)\cdot\tilde{T}_t]$$

As a consequence, replacing the tangent vector T_t by its conditional expectation $\mathbb{E}[T_s|X_s]$ does not affect the value of the expectation. It is thus possible during the numerical simulation to merge close particles and associate the obtained particle with a tangent vector equal to the mean of the merged particles.

6.4.4 Numerics

The particle method described in section 6.4.1 has been studied numerically on the following toy model. Let $(V_\lambda)_{\lambda\in\mathbb{R}}$ be the following family of potentials

$$V_\lambda(x) = \mathbf{1}_{|x|>1}(x - sgn(x))^2 + \mathbf{1}_{|x|\leq 1}\lambda(x^2-1)^2,$$

corresponding to a potential quadratic at infinity with a double well. The parameter λ describes the height of the potential barrier between the two wells. The graph of these functions is plotted on Figure 6.1, for $lambda = 0.8$ (lower curve), $lambda = 1$ (middle curve) and $lambda = 1.2$ (upper curve).

On Figure 6.2, an approximation of

$$\partial_\lambda\left(\frac{\int_{\mathbb{R}} x^2 e^{-V_\lambda(x)}\mathrm{d}x}{\int_{\mathbb{R}} e^{-V_\lambda(x)}\mathrm{d}x}\right) \tag{6.23}$$

at $\lambda = 1$ has been computed with two different methods :

- on the left, through a Monte Carlo method with 20000 independent realizations of a Langevin process on the time interval $[0, 10]$, relying on the ergodicity of the Langevin dynamics with respect to the measure e^{-V_λ} ;

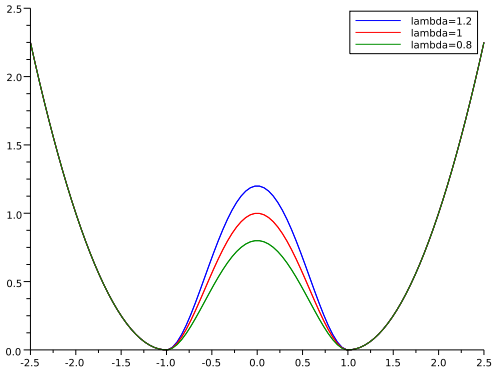

Fig. 6.1. Graph of the potential V_λ for λ with respective values 0.8, 1, and 1.2.

– on the right, the with the resampling procedure described in 6.4.1, also with 20000 particles. There have been four resampling at regular steps on the interval $[0, 10]$.

The exact value of the derivative 6.23 has been plotted on Figure 6.2 as an horizontal line. The curve is the approximate value computed in real time, with a 95% confidence interval.

The confidence intervals have been computed using 400 independent realizations of the described algorithm. At final time $t = 10$, the variance is 0.0155 without resampling, and 0.0038 with resampling.

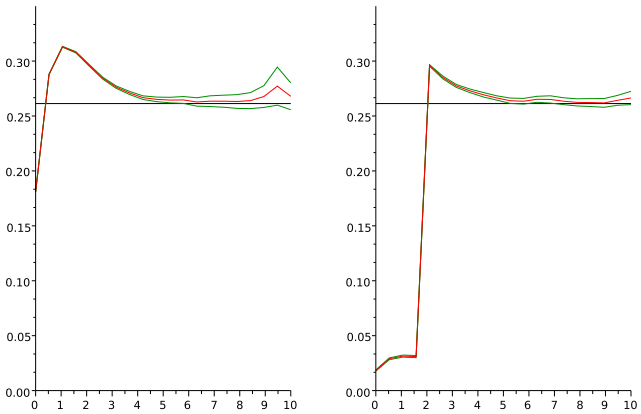

Fig. 6.2. Approximation of 6.23. Left : using a Monte Carlo procedure on a Langevin dynamics. Right : using resampling.

Partie IV

Interprétation probabiliste d'une équation hyperbolique

7

Convergence d'une approximation particulaire pour une loi de conservation scalaire fractionnaire

Résumé : Dans cette partie, nous nous intéressons à l'approximation de lois de conservations scalaires par des systèmes de particules en interaction probabiliste. Ces lois de conservations pourront inclure un terme de diffusion d'ordre non entier $\alpha \in (0, 2]$. Quand $\alpha \leq 1$ ou si le terme de diffusion est absent (équation non visqueuse), la solution de la loi de conservation se caractérise par une formulation entropique. L'interprétation probabiliste de la loi de conservation fractionnaire est basée sur une équation différentielle stochastique dirigée par un processus de Lévy $\alpha-$stable et comportant un terme de dérive non-linéaire au sens de McKean. L'approximation particulaire correspondante est construite en discrétisant l'équation en temps grâce à un schéma d'Euler et en remplaçant la non linéarité par une interaction entre les particules. À chaque particule est affecté un signe dépendant de sa condition initiale. À chaque pas de discrétisation, on tue les éventuels couples de particules suffisamment proches ayant des signes opposés, puisque la contribution des croisements de tels couples de particules aurait le mauvais signe au vu de la formulation entropique. Nous prouvons la convergence de l'approximation particulaire de la loi de conservation quand le nombre de particules tend vers l'infini, alors que la distance de meurtre, le pas de temps, et dans le cas non-visqueux, le coefficient de diffusion tendent vers 0 dans des proportions précises dépendant de la position relative de α et de sa valeur critique 1.

Mots-Clés : Equations aux dérivées partielles non-linéaires, systèmes de particules en interaction, schéma d'Euler, processus de Lévy α-stables.

Abstract : In this chapter, we are interested in approximating the solution to scalar conservation laws using systems of interacting stochastic particles. The scalar conservation law may involve a fractional Laplacian term of order $\alpha \in (0, 2]$. When $\alpha \leq 1$ as well as in the absence of this term (inviscid case), its solution is characterized by entropic inequalities. The probabilistic interpretation of the scalar conservation law is based on a stochastic differential equation driven by an α-stable process and involving a drift nonlinear in the sense of McKean. The particle system is constructed by discretizing this equation in time by the Euler scheme and replacing the nonlinearity by interaction. Each particle carries a signed weight depending on its initial position. At each discretization time we kill the couples of particles with opposite weights and positions closer than a threshold since the contribution of the crossings of such particles has the wrong sign in the derivation of the entropic inequalities. We prove convergence of the particle approximation to the solution of the conservation law as the number of particles tends to infinity whereas the discretization step, the killing threshold and, in the inviscid case, the coefficient multiplying the stable increments tend to 0 in some precise asymptotics depending on whether α is larger than the critical level 1.

Keywords : Nonlinear partial differential equations, interacting particle systems, Euler scheme, *alpha*-stable Lévy processes.

Introduction

We are interested in providing a numerical probabilistic scheme for the fractional scalar conservation law of order α

$$\partial_t v(t,x) + \sigma^\alpha(-\Delta)^{\frac{\alpha}{2}} v(t,x) + \partial_x A(v(t,x)) = 0, \quad (t,x) \in \mathbb{R}_+ \times \mathbb{R}, \tag{7.1}$$

where $-(-\Delta)^{\frac{\alpha}{2}}$ is the fractional Laplacian operator of order $0 < \alpha \leq 2$ (defined in Section 2), and A is a function of class $\mathcal{C}^1$ from $\mathbb{R}$ to $\mathbb{R}$. We also consider the equation obtained by letting $\sigma \to 0$ in (7.2), namely the inviscid conservation law

$$\partial_t v(t,x) + \partial_x A(v(t,x)) = 0, \quad (t,x) \in \mathbb{R}_+ \times \mathbb{R}. \tag{7.2}$$

These equations have already been studied intensively from a deterministic point of view, see for example [5,12,13,22,23,33] and references therein.

In [36, 39], these equations are interpreted as Fokker-Planck equations associated to some stochastic differential equations nonlinear in the sense of McKean, which can be approximated by a particle system. Interacting particle systems have already been used for the study of general nonlinear Markov semigroups in [42, 43]. However, in our setting, the dependence of the drift in the law of the solution is not regular enough to directly apply those results.

We introduce an Euler time discretization of the particle system and show the convergence of its empirical cumulative distribution function to the solution of (7.1). We also study its convergence to the solution of (7.2) as the parameter σ goes to 0.

Euler schemes for viscous conservation laws have already been studied in [14], [15], [17] or [18], where a convergence rate of $1/\sqrt{N} + \sqrt{\Delta t}$ is derived in the case $\alpha = 2$, N denoting the number of particles, and Δt being the time step.

To give the probabilistic interpretation to (7.1) we consider the space derivative $u = \partial_x v$ of a solution v to equation (7.1), which formally satisfies

$$\partial_t u_t = -\sigma^\alpha(-\Delta)^{\frac{\alpha}{2}} u_t - \partial_x \left(A'(H * u_t)u_t\right), \tag{7.3}$$

where $H = \mathbf{1}_{[0,\infty)}$ denotes the Heaviside function. When u_0 is a probability measure, that is, when the initial condition v_0 of Equation (7.1) is a cumulative distribution function, Equation (7.3) is the Fokker-Planck equation associated to the following nonlinear stochastic differential equation

$$\begin{cases} \mathrm{d}X_t &= \sigma \mathrm{d}L_t^\alpha + A'(H * u_t(X_t))\mathrm{d}t \\ u_t &= \text{ law of } X_t \end{cases},$$

where L_t^α is a Markov process with generator $-(-\Delta)^{\frac{\alpha}{2}}$, namely $\sqrt{2}$ times a Brownian motion for $\alpha = 2$, and a stable Lévy process with index α in the case $\alpha < 2$, that is to say a pure jump Lévy process whose Lévy measure is given by $c_\alpha \mathrm{d}y/|y|^{1+\alpha}$, where c_α is some positive constant.

We can still give a probabilistic interpretation to Equation (7.1) if the initial condition v_0 has bounded variation, is right continuous and not constant. Indeed, in that case v_0 can be written as $v_0(x) = a + \int_{-\infty}^x \mathrm{d}u_0(y) = a + H * u_0(x)$ for some finite measure u_0. By replacing $v_0(x)$ by $(v_0(x) - a)\left(|u_0|(\mathbb{R})\right)^{-1}$ and $A(x)$ by $A(a + x|u_0|(\mathbb{R}))(|u_0|(\mathbb{R}))^{-1}$ in (7.1) ($|u_0|$ denoting the total variation of the measure u_0,), one can assume without loss of generality that $a = 0$ and that $|u_0|$ is a probability measure. We denote by $\gamma = \mathrm{d}u_0/\mathrm{d}|u_0|$ the Radon-Nikodym density of u_0 with respect to its total variation. Notice that γ takes values in $\{\pm 1\}$.

Then, Equation (7.3) is the Fokker-Planck equation associated to

$$\begin{cases} \mathrm{d}X_t &= \sigma \mathrm{d}L_t^\alpha + A'(H * \tilde{P}_t(X_t))\mathrm{d}t \\ P &= \text{ law of } X \end{cases}, \tag{7.4}$$

where $\tilde{P}$ denotes the measure defined on the Skorokhod space $\mathcal{D}$ of càdlàg functions from $[0,\infty)$ to $\mathbb{R}$ by its Radon-Nikodym density $\mathrm{d}\tilde{P}/\mathrm{d}P = \gamma(f(0))$, with f the canonical process on $\mathcal{D}$, and $\tilde{P}_t$ denotes its time marginal at time t, *i.e* the measure defined by $\tilde{P}_t(B) = \int_{\mathcal{D}} \gamma(f(0))\mathbf{1}_B(f(t))\mathrm{d}P(f)$, for any B in the Borel $\sigma-$field of $\mathbb{R}$.

The rest of the paper is organized as follows :
In Section 1 we define the particle approximation for the stochastic differential equation (7.4).
Section 2 is devoted to the definition of the different notions of solutions used in the article.
In Section 3, we analyze the convergence of the time-discretized particle system to the solution of the conservation law in different settings : for both a constant or vanishing diffusion coefficient and any value of $0 < \alpha \leq 2$.
Finally, we present some numerical simulations in Section 4. Those simulations are compared with the results of a deterministic method described in [29].

In the following, the letter K denotes some positive constant whose value can change from line to line.

7.1 The particle approximation

In this section we construct a discretization of (7.4) consisting of both a particle approximation in order to approximate the law of the solution and an Euler discretization to make the particles evolve in time. The idea is to introduce N particles $X^{N,1},\dots,X^{N,N}$ which are N interacting copies of the stochastic differential equation (7.4), where the actual law P of the process is replaced by the empirical distribution of the particles $N^{-1}\sum_{i=1}^N \delta_{X^{N,i}}$.

In continuous time, those particles are driven by N independent Brownian motions or stable Lévy processes with index α and undergo a drift given by $A'(H * \tilde{\mu}_t^N(.))$, with $\tilde{\mu}_t^N = N^{-1}\sum_{i=1}^N \gamma(X_0^{N,i})\delta_{X_t^{N,i}}$. The natural way to introduce the measure $\tilde{\mu}_t^N$ in the dynamics is to give each particle a signed weight equal to the evaluation of γ at the initial position of the particle. Then, $H * \tilde{\mu}_t^N(x)$ is simply given by the sum of weights of particles situated left from x.

The entropy solution to (7.1) has a non-increasing total variation (see [3]), which can be interpreted probabilistically as a compensation of merging sample paths having opposite signs. For a more precise statement in the case $\alpha = 2$, see Lemma 2.1 in [36]. It is thus natural to adapt this behavior in our particle approximation by killing any merging couple of particles with opposite signs.

In [36] Jourdain proves, for $\alpha = 2$ in continuous time, the convergence of the particle system to the solution of the nonlinear stochastic differential equation through a propagation-of-chaos result. Moreover, the convergence of the signed cumulative distribution function $H * \tilde{\mu}_t^N$ to the solution to Equation (7.1) is also proved, as well as convergence to the solution to the inviscid equation as $\sigma \to 0$. In [39] the same results are generalized to the case $1 < \alpha < 2$, assuming $\gamma = 1$ in the case of a vanishing viscosity. However, the existence of both the nonlinear process and the particle system is a much more challenging problem in the case $\alpha \leq 1$, since the driving Lévy process is somehow weaker than the drift. This remains, to our knowledge, an open question. Even recent papers treating stochastic differential equations including a drift term only deal with the case $\alpha > 1$, see for example [57, 73].

A natural way to ensure existence of the approximation is to transpose the problem in discrete time using an Euler discretization. In discrete time, the probability of seeing two particles actually merging is 0. To adapt the murders from the continuous time setting, we thus kill, at each time step, any couple of particles with opposite signs separated by a distance smaller than a given threshold ε_N going to zero as N goes to ∞. Though, one has to be careful, since one can have more than two particles lying in a small interval of length ε_N. Precisely, the particles are killed in the following way : kill the leftmost couple of particles at consecutive positions separated by a distance smaller than the threshold ε_N and with opposite signs. Then, recursively apply the same

algorithm to the remaining particles. This can be done with a computational cost of order $\mathcal{O}(N)$. The essential properties satisfied by this killing procedure are the following :

- to each killed particle is attached another killed particle, which has opposite signs and lies at a distance at most ε_N of the first particle.
- after the killing there is no couple of particles with opposite signs in a distance smaller than ε_N.
- the exchangeability of the particles is preserved.
- after the murder, the quantity $H * \tilde{\mu}_t^N(X_t^{N,i})$ remains the same for any surviving particle.

We are going to describe the killed processes by a couple (f, κ) in the space $\mathcal{K} = \mathcal{D} \times [0, \infty]$ of càdlàg functions f from $[0, \infty)$ to $\mathbb{R}$ endowed with a death time $\kappa \in [0, \infty]$. The space $\mathcal{K}$ is endowed with the product metric $d((f, \kappa_f), (g, \kappa_g)) = d_S(f, g) + |\arctan(\kappa_f) - \arctan(\kappa_g)|$, where d_S is the Skorokhod metric on $\mathcal{D}$, so that $(\mathcal{K}, d)$ is a complete metric space. It could seem more natural to consider the space $\mathcal{D}([0, \infty), \mathbb{R} \cup \{\partial\})$ of paths taking values in $\mathbb{R}$ endowed with a cemetery point ∂. However the corresponding topology is too strong to prove Proposition 7.3.4.

The precise description of the process is the following : each particle will be represented by a couple $(X^{N,i}, \kappa_i^N) \in \mathcal{K}$. Let $(X_0^i)_{i \in \mathbb{N}}$ be a sequence of independent random variables with common distribution $|u_0|$ and let $h_N > 0$ denote the time step of the Euler scheme. At time 0, kill the particles according to the preceding rules, that is to say, set $\kappa_i^N = 0$ for killed particles, which will not move anymore. Those particles will not be taken into account anymore. Now, by induction, suppose that the particle system has been defined up to time kh_N, and kill the particles according to the preceding rules (*i.e.* set $\kappa_i^N = kh_N$ and $X_t^{N,i} = X_{kh_N}^{N,i}$ for all $t \geq kh_N$, if the particle with index i is one of those). Then let the particles still alive evolve up to time $(k+1)h_N$ according to

$$\mathrm{d}X_t^{N,i} = A'\left(\frac{1}{N} \sum_{\kappa_j^N > kh_N} \gamma(X_0^j) \mathbf{1}_{X_{kh_N}^{N,j} \leq X_{kh_N}^{N,i}}\right) \mathrm{d}t + \sigma_N \mathrm{d}L_t^i,$$

where $(L^i)_{i \in \mathbb{N}}$ is a sequence of independent α-stable Lévy processes for $\alpha < 2$, or a sequence of independent copies of $\sqrt{2}$ times Brownian motion, which are independent of the sequence $(X_0^i)_{i \in \mathbb{N}}$. The particle system is thus well-defined, by induction.

Let $\mu^N = N^{-1} \sum_{i=1}^N \delta_{(X^{N,i}, \kappa_i^N)} \in \mathcal{P}(\mathcal{K})$ be the empirical distribution of the particles. For a probability measure Q on $\mathcal{K}$ and $t \geq 0$, we define a signed measure $\tilde{Q}_t$ on $\mathbb{R}$ by :

$$\tilde{Q}_t(B) = \int_{\mathcal{K}} \mathbf{1}_B(f(t)) \mathbf{1}_{\kappa > t} \gamma(f(0)) \mathrm{d}Q(f, \kappa),$$

for any B in the Borel $\sigma-$field of $\mathbb{R}$. With these notations, on the interval $[kh_N, (k+1)h_N)$, a particle, provided it is still alive, satisfies

$$\mathrm{d}X_t^{N,i} = A'\left(H * \tilde{\mu}_{kh_N}^N\left(X_{kh_N}^{N,i}\right)\right) \mathrm{d}t + \sigma_N \mathrm{d}L_t^i.$$

Notice that the sum of the weights of alive particles $\tilde{\mu}_t^N(\mathbb{R}) = N^{-1} \sum_{\kappa_i^N > t} \gamma(X_0^i)$ is constant in time, since the particles are killed by couples of opposite signs.

7.2 Notion of solutions

In this section, we recall the different notions of solutions that are associated to the equations (7.1) and (7.2). Indeed, due to the shock-creating term $\partial_x(A(u_t))$, the notion of weak solution is too weak, and does not provide uniqueness when the diffusion term is not regularizing enough. The best suited notion in those cases is the notion of entropy solution.

In [44], Kruzhkov shows that for $v_0 \in \mathbb{L}^\infty((0,\infty))$ existence and uniqueness hold for entropy solutions to (7.2), defined as functions $v \in \mathbb{L}^\infty((0,\infty)\times\mathbb{R})$ satisfying, for any smooth convex function η, any nonnegative smooth function g with compact support on $[0,\infty)\times\mathbb{R}$ and any ψ satisfying $\psi' = \eta' A'$, the entropic inequality

$$\int_{\mathbb{R}} \eta(v_0) g_0 + \int_0^\infty \left(\int_{\mathbb{R}} \eta(v_t)\partial_t g_t + \psi(v_t)\partial_x g_t \right) \mathrm{d}t \geq 0. \tag{7.5}$$

It is well known that this entropy solution can be obtained as the limit of weak solutions to (7.1) as $\sigma \to 0$ in the case $\alpha = 2$.

Weak solutions to (7.1) (see [36]) are defined as functions $v \in \mathbb{L}^\infty((0,\infty)\times\mathbb{R})$ satisfying, for all smooth functions g with compact support in $[0,\infty)\times\mathbb{R}$,

$$\int_{\mathbb{R}} v_0 g_0 + \int_0^\infty \int_{\mathbb{R}} v_t \partial_t g_t \mathrm{d}t - \sigma^\alpha \int_0^\infty \int_{\mathbb{R}} v_t (-\Delta)^{\frac{\alpha}{2}} g_t \mathrm{d}t + \int_0^\infty \int_{\mathbb{R}} A(v_t)\partial_x g_t \mathrm{d}t = 0. \tag{7.6}$$

For $\alpha < 2$, we denote by $(-\Delta)^{\frac{\alpha}{2}}$ the fractional symmetric differential operator of order α, that can be defined through the Fourier transform :

$$\widehat{(-\Delta)^{\frac{\alpha}{2}} u}(\xi) = |\xi|^\alpha \hat{u}(\xi).$$

An equivalent definition for $(-\Delta)^{\frac{\alpha}{2}}$ uses an integral representation

$$(-\Delta)^{\frac{\alpha}{2}} u(x) = c_\alpha \int_{\mathbb{R}} \frac{u(x+y) - u(x) - \mathbf{1}_{|y|\leq r} u'(x) y}{|y|^{1+\alpha}} \mathrm{d}y$$

for any $r \in (0,\infty)$ and some fixed constant c_α (see [31]), depending on the definition of the Fourier transform.

In [36] and [39], it has been proven, using probabilistic arguments, that existence and uniqueness hold for weak solutions of (7.1), for $1 < \alpha \leq 2$. Similar results had already been proven in [30] using analytic arguments. However, for $0 < \alpha \leq 1$, the diffusive term of order α in (7.1) is somehow dominated by the shock-creating term, which is of order 1, so that a weak formulation does not ensure uniqueness for the solution. We thus have to strengthen the notion of solution, and use entropy solutions to (7.1), defined in [3] as functions v in $\mathbb{L}^\infty((0,\infty)\times\mathbb{R})$ satisfying the relation

$$\begin{aligned} &\int_0^\infty \eta(v_0) g_0 + \int_0^\infty \int_{\mathbb{R}} \left(\eta(v_t)\partial_t g_t + \psi_t(v_t)\partial_x g_t\right) \mathrm{d}t \\ &\quad + c_\alpha \int_0^\infty \int_{\mathbb{R}} \int_{\{|y|>r\}} \eta'(v_t(x)) \frac{v_t(x+\sigma y) - v_t(x)}{|y|^{1+\alpha}} g_t(x) \mathrm{d}y\mathrm{d}x\mathrm{d}t \\ &\quad + c_\alpha \int_0^\infty \int_{\mathbb{R}} \int_{\{|y|\leq r\}} \eta(v_t(x)) \frac{g_t(x+\sigma y) - g_t(x) - \sigma y \partial_x g_t(x)}{|y|^{1+\alpha}} \mathrm{d}y\mathrm{d}x\mathrm{d}t \geq 0 \end{aligned} \tag{7.7}$$

for any $r > 0$, any nonnegative smooth function g with compact support in $[0,\infty)\times\mathbb{R}$, any smooth convex function $\eta : \mathbb{R} \to \mathbb{R}$ and any ψ satisfying $\psi' = \eta' A'$. Notice that from the convexity of η, the entropic formulation (7.7) for a parameter r implies the entropic formulation with parameter $r' > r$. Also notice, using the functions $\eta(x) = \pm x$ that an entropy solution to (7.1) is a weak solution to (7.1).

In [3], Alibaud shows that existence and uniqueness hold for entropy solutions of (7.1) provided that the initial condition v_0 lies in $\mathbb{L}^\infty(\mathbb{R})$. The entropy solution then lies in the space $\mathcal{C}\left([0,\infty), \mathbb{L}^1(\mathrm{d}x/(1+x^2))\right)$. He also proves that the entropy solution to (7.1) converges to the entropy solution to (7.2) in the space $\mathcal{C}([0,T], \mathbb{L}^1_{\mathrm{loc}}(\mathbb{R}))$ as $\sigma \to 0$.

7.3 Statement of the results

The aim of this article is to prove the three following convergence result, each one corresponding to a particular setting.

Theorem 7.3.1 *Assume $0 < \alpha \leq 1$. Let $\sigma_N \equiv \sigma$ be a constant sequence. Let ε_N and h_N be two sequences going to zero and satisfying the inequalities*

$$N^{-\lambda} \leq 4 \sup_{[-1,1]} |A'| h_N \leq \varepsilon_N, \text{ and } N^{-1/\alpha} \leq N^{-1/\lambda} \varepsilon_N$$

for some positive λ. For $\alpha = 1$, also assume $h_N \leq \varepsilon_N N^{-1/\lambda}$. Then, for any $T > 0$,

$$\lim_{N\to\infty} \int_0^T \mathbb{E} \left\| H * \tilde{\mu}_t^N - v_t \right\|_{\mathbb{L}^1\left(\frac{\mathrm{d}x}{1+x^2}\right)} \mathrm{d}t = 0,$$

where v_t denotes the entropy solution to the fractional conservation law (7.1).

Theorem 7.3.2 *Let ε_N, h_N and σ_N be three sequences going to zero such that*

$$N^{-\lambda} \leq 4 \sup_{[-1,1]} |A'| h_N \leq \varepsilon_N$$

for some $\lambda > 0$. If $\alpha > 1$, also assume $\sigma_N \leq \varepsilon_N^{1-\frac{1}{\alpha}} N^{-\frac{1}{\lambda}}$. Then, for any $T > 0$,

$$\lim_{N\to\infty} \int_0^T \mathbb{E} \left\| H * \tilde{\mu}_t^N - v_t \right\|_{\mathbb{L}^1\left(\frac{\mathrm{d}x}{1+x^2}\right)} \mathrm{d}t = 0,$$

where v_t denotes the entropy solution to the inviscid conservation law (7.2).

The additional assumption for $\alpha > 1$ comes from the fact that in this case, the dominant term is the diffusion, while in the limit there is no diffusion anymore. The assumption ensures that the diffusion is weak enough not to perturb the approximation. For $\alpha \leq 1$, the dominant term is the drift, as in the limit, so that no additional condition is needed.

Theorem 7.3.3 *Assume $1 < \alpha \leq 2$. Let $\sigma_N \equiv \sigma$ be a constant sequence, and let ε_N and h_N be two sequences going to zero. Then, for any $T > 0$,*

$$\lim_{N\to\infty} \int_0^T \mathbb{E} \left\| H * \tilde{\mu}_t^N - v_t \right\|_{\mathbb{L}^1\left(\frac{\mathrm{d}x}{1+x^2}\right)} \mathrm{d}t = 0,$$

where v_t denotes the weak solution to the fractional conservation law (7.1).

In order to prove those three theorems, we will have to control the probability of seeing particles merging. In the case $\alpha < 2$, this is mainly due to the conjunction of the small jumps of the stable process and the drift coefficient, while the large jumps of the stable term do not play an essential role. As a consequence, for $\alpha < 2$, we consider another family of evolutions coinciding with the Euler scheme on the time discretization grid, for which we consider differently the jumps which are smaller or larger than a given threshold r. The choice of this parameter has to be linked to the parameter r appearing in the entropic formulation (7.7), since they play a similar role : the third term in (7.7) corresponds to the effect of jumps larger than r in the driving Lévy process and the fourth term corresponds to jumps smaller than r. This evolution is designed so that on the first half of each time step, the process will evolve according to the drift and the small jumps, and on the second half of each time step, it will evolve according to the large jumps. More precisely, let

$$\nu^i(\mathrm{d}y, \mathrm{d}t) = \sum_{\Delta L_t^i \neq 0} \delta_{(\Delta L_t^i, t)}$$

be the jump measure associated to the Lévy process L^i and let

$$\tilde{\nu}^i(\mathrm{d}y, \mathrm{d}t) = \nu^i(\mathrm{d}y, \mathrm{d}t) - c_\alpha \frac{\mathrm{d}y\mathrm{d}t}{|y|^{1+\alpha}}$$

be the corresponding compensated measure, so that

$$L_t^i = \int_{(0,t]\times\{|y|>r\}} y\nu^i(\mathrm{d}y, \mathrm{d}t) + \int_{(0,t]\times\{|y|\leq r\}} y\tilde{\nu}^i(\mathrm{d}y, \mathrm{d}t),$$

where the right hand side does not depend on r. We define the process $X^{N,i,r}$ by

$$X^{N,i,r} = X_0^i + \sigma_N L^{N,i,r} + \sigma_N \Lambda^{N,i,r} + \mathcal{A}^{N,i},$$

where

- $L_t^{N,i,r}$ is the large jumps part defined by

$$L_t^{N,i,r} = \int_{(0,a(t)]\times\{|y|>r\}} y\nu^i(\mathrm{d}y, \mathrm{d}s),$$

  where $a(t) = \begin{cases} kh_N & \text{for } t \in [kh_N, (k+1/2)h_N] \\ kh_N + 2(t-(k+1/2)h_N) & \text{for } t \in [(k+1/2)h_N, (k+1)h_N] \end{cases}$. This process is constant on intervals $[kh_N, (k+1/2)h_N]$ and behaves like a Lévy process with jump measure $\mathbf{1}_{|y|>r} 2c_\alpha \mathrm{d}y/|y|^{1+\alpha}$ on intervals $[(k+1/2)h_N, (k+1)h_N]$.
- $\Lambda_t^{N,i,r}$ is the small jumps part, defined by

$$\Lambda_t^{N,i,r} = \int_{(0,b(t)]\times\{|y|\leq r\}} \tilde{\nu}^i(\mathrm{d}y, \mathrm{d}s),$$

  where $b(t) = \begin{cases} kh_N + 2(t-kh_N) & \text{for } t \in [kh_N, (k+1/2)h_N] \\ (k+1)h_N & \text{for } t \in [(k+1/2)h_N, (k+1)h_N] \end{cases}$. This term behaves like a Lévy process with jump measure $\mathbf{1}_{|y|\leq r} 2c_\alpha \mathrm{d}y/|y|^{1+\alpha}$ on intervals $[kh_N, (k+1/2)h_N]$ and is constant on intervals $[(k+1/2)h_N, (k+1)h_N]$. Notice that the process $\Lambda^{N,i,r}$ is a martingale.
- $\mathcal{A}^{N,i}$ is the drift part, which satisfies $\mathcal{A}_0^{N,i} = 0$, is constant over each interval $[(k+1/2)h_N, (k+1)h_N]$, and evolves as a piecewise affine process with derivative $2A'(H * \tilde{\mu}_{kh_N}^N(X_{kh_N}^{N,i}))$ on intervals $[kh_N, (k+1/2)h_N]$.

One can check that for any r, the process $(X^{N,1,r}, \ldots, X^{N,N,r})$ is equal to $(X^{N,1}, \ldots, X^{N,N})$ on the time discretization grid up to killing time. Conditionally on the positions of the particles at time kh_N, the particles evolve independently on $[kh_N, (k+1)h_N]$, and the evolution on $[kh_N, (k+1/2)h_N]$ is independent of the evolution on $[(k+1/2)h_N, (k+1)h_N]$. Since the entropic formulation (7.7) with parameter r is stronger than the one with parameter $r' \geq r$, we have to make the parameter r tend to zero in order to prove the entropic formulation for any parameter. However, this convergence has to satisfy some conditions with respect to N, h_N and ε_N. We will explain later why a suitable sequence r_N exists under the conditions given in the statement of Theorem 7.3.1.

In order to prove Theorems 7.3.1 and 7.3.2, we introduce $\mu^{N,r}$ the empirical distribution of the processes $(X^{N,i,r}, \kappa_i^N)$:

$$\mu^{N,r} = \frac{1}{N}\sum_{i=1}^N \delta_{(X^{N,i,r}, \kappa_i^N)} \in \mathcal{P}(\mathcal{K}),$$

and by $\pi^{N,r}$ the law of $\mu^{N,r}$.

The following proposition is the first step in the proof of Theorems 7.3.1, 7.3.2 and 7.3.3.

Proposition 7.3.4 – *Assume $\alpha < 2$. For any bounded sequences (h_N), (σ_N) and (ε_N), and for any sequence (r_N), the family of probability measures $(\pi^{N,r_N})_{N\in\mathbb{N}}$ is tight in $\mathcal{P}(\mathcal{P}(\mathcal{K}))$.*
– *Denote by π^N the law of μ^N. For any bounded sequences (h_N), (σ_N) and (ε_N), the family of probability measures $(\pi^N)_{N\in\mathbb{N}}$ is tight in $\mathcal{P}(\mathcal{P}(\mathcal{K}))$.*

Proof. We first check the tightness of the family $(\pi^{N,r_N})_{N\in\mathbb{N}}$.

As stated in [64], checking the tightness of the sequence π^{N,r_N} boils down to checking the tightness of the sequence $(\mathrm{Law}(X^{N,1,r_N}, \kappa_1^N))$. Owing to the product-space structure, we can check tightness for $X^{N,1,r_N}$ and κ_1^N separately.

Of course, tightness for κ_1^N is straightforward since it lies on the compact space $[0,\infty]$, and it is enough to check tightness for the laws of the path $(X^{N,1,r_N})$. For simplicity, we will assume that $A = 0$, which is not restrictive since A' is a bounded function so that the perturbation induced by A belongs to a compact subset of the space of continuous functions, from Ascoli's theorem (also notice that the addition functional from $\mathcal{D}\times\mathcal{C}([0,\infty))$ to $\mathcal{D}$ is continuous). We use Aldous' criterion to prove tightness (see [2]). First, the sequences $(X_0^{N,1,r_N})_{N\in\mathbb{N}}$ and $(\sup_{[0,T]}|\Delta X^{N,1,r_N}|)_{N\in\mathbb{N}}$ are tight, since $(X_0^{N,1,r_N})$ is constant in law and $\left(\sup_{[0,T]}|\Delta X^{N,1,r_N}|\right)_N$ is dominated by the identically distributed sequence

$$\left(\left(\sup_N \sigma_N\right) \sup_{[0,T+\sup_N h_N]} \left|\Delta L^1\right|\right)_N .$$

Then let τ_N be a stopping time of the natural filtration of $X^{N,1,r_N}$ taking finitely many values, and let $(\delta_N)_{N\in\mathbb{N}}$ be a sequence of positive numbers going to 0 as $N\to\infty$. One can write

$$\begin{aligned}\mathbb{P}\left(\left|X^{N,1,r_N}_{\tau_N+\delta_N} - X^{N,1,r_N}_{\tau_N}\right| \geq \varepsilon\right) \leq& \mathbb{P}\left(\sigma_N\left|\Lambda^{N,1,r_N}_{\tau_N+\delta_N} - \Lambda^{N,1,r_N}_{\tau_N}\right| \geq \varepsilon/2\right)\\ &+ \mathbb{P}\left(\sigma_N\left|L^{N,1,r_N}_{\tau_N+\delta_N} - L^{N,1,r_N}_{\tau_N}\right| \geq \varepsilon/2\right) \qquad (7.8)\\ \leq& \mathbb{P}\left(\sup_{t\in[0,\delta_N]} \sigma_N|\mathcal{L}_t^{\leq r_N}| \geq \varepsilon/2\right) + \mathbb{P}\left(\sup_{t\in[0,\delta_N]} \sigma_N|\mathcal{L}_t^{>r_N}| \geq \varepsilon/2\right),\end{aligned}$$

where

$$\mathcal{L}_t^{\leq r} = \int_{(0,t]\times\{|y|\leq r\}} y\tilde{\nu}(\mathrm{d}y,\mathrm{d}t) \quad \text{and} \quad \mathcal{L}_t^{>r} = \int_{(0,t]\times\{|y|>r\}} y\nu(\mathrm{d}y,\mathrm{d}t),$$

the measure ν being the jump measure of some Lévy process $\mathcal{L}$ with Lévy measure $2c_\alpha \mathrm{d}y/|y|^{1+\alpha}$, and $\tilde{\nu}$ is the compensated measure of ν. Now, using the maximal inequality for the martingale $(\mathcal{L}_t^{\leq r_N})_{t\in[0,\delta_N]}$, noticing that $(\mathcal{L}_{\delta_N}^{\leq r})_{r\in[0,1]}$ is also a martingale, we deduce

$$\begin{aligned}\mathbb{P}\left(\sup_{t\in[0,\delta_N]} |\mathcal{L}_t^{\leq r_N}| \geq \varepsilon/2\sigma_N\right) &\leq \sup_{r\in[0,\sup_N r_N]} \mathbb{P}\left(\sup_{t\in[0,\delta_N]} |\mathcal{L}_t^{\leq r}| \geq \varepsilon/2\sigma_N\right)\\ &\leq 2\sigma_N\varepsilon^{-1} \sup_{r\in[0,\sup_N r_N]} \mathbb{E}\left(|\mathcal{L}_{\delta_N}^{\leq r}|\right)\\ &= 2\sigma_N\varepsilon^{-1}\mathbb{E}\left(|\mathcal{L}_{\delta_N}^{\leq \sup_N r_N}|\right)\\ &\underset{N\to\infty}{\longrightarrow} 0.\end{aligned}$$

For the large jumps parts, one writes,

$$\begin{aligned}\mathbb{P}\left(\sup_{t\in[0,\delta_N]}|\mathcal{L}_t^{>r_N}|\geq\varepsilon/2\sigma_N\right)&\leq\mathbb{P}\left(\sup_{t\in[0,\delta_N]}|\mathcal{L}_t|+\sup_{t\in[0,\delta_N]}|\mathcal{L}_t^{\leq r_N}|\geq\varepsilon/2\sigma_N\right)\\&\leq\mathbb{P}\left(\sup_{t\in[0,\delta_N]}|\mathcal{L}_t|\geq\varepsilon/4\sigma_N\right)+\mathbb{P}\left(\sup_{t\in[0,\delta_N]}|\mathcal{L}_t^{\leq r_N}|\geq\varepsilon/4\sigma_N\right)\\&\underset{N\to\infty}{\to}0.\end{aligned}$$

As a consequence, the family $(\mathrm{Law}(X^{N,1,r_N}))_{N\in\mathbb{N}}$ is tight in $\mathcal{D}$.

Thus, the family $(\pi^{N,r_N})_{N\in\mathbb{N}}$ is tight.

The proof is essentially the same for the tightness of $(\pi^N)_{N\in\mathbb{N}}$, with a few simplifications, since we do not treat separately large and small jumps. It also adapts in the case $\alpha=2$, since the Gaussian distribution has thinner tails than the $\alpha-$stable distribution for $\alpha<2$.

The use of the path space $\mathcal{K}$ instead of $\mathcal{D}([0,\infty),\mathbb{R}\cup\{\partial\})$ for a cemetery point ∂ is crucial in the proof of Proposition 7.3.4, since in the latter case, we need to control the jumps occuring close to the death time in order to prove tightness. The following example is illustrative : if we consider a sequence f_n of paths starting at 0, jumping to 1 at time $1-1/n$, and being killed at time 1, then f_n does not converge in $\mathcal{D}([0,\infty),\mathbb{R}\cup\{\partial\})$, while it does in $\mathcal{K}$.

The following lemma deals with the initial condition of the particle system.

Lemma 7.3.5 *If π^∞ is the limit of some subsequence of π^N or π^{N,r_N}, then for π^∞-almost all Q, for all A in the Borel $\sigma-$field of $\mathbb{R}$,*

$$Q_0(A):=\int_{\mathbb{R}}\mathbf{1}_{\kappa>0}\mathbf{1}_{f(0)\in A}\mathrm{d}Q(f,\kappa)=|u_0|(A). \tag{7.9}$$

In particular, κ is $Q-$almost surely positive for π^∞-almost all Q.

Proof. In a first time, we control the probability of seeing a partincle dying within a short time.

Let us write the Hahn decomposition $u_0^+-u_0^-$ of the measure u_0, the measures u_0^+ and u_0^- being positive measures supported by two disjoint sets B^+ and B^-. From the inner regularity of the measure u_0^+, for any $\delta>0$, one can find a closed set $F^+\subset B^+$ such that $u_0^+(F^+)\geq u_0^+(B^+)-\delta$. The complement set $O^-=(F^+)^c$ is then an open subset of $\mathbb{R}$, which can thus be decomposed as a countable union of disjoint open intervals $O^-=\bigcup_{m=1}^\infty]a_m,b_m[$. For a large enough integer M, and for $\varepsilon_\delta>0$ small enough, the set $O^\delta=\bigcup_{m=1}^M]a_m+\varepsilon_\delta,b_m-\varepsilon_\delta[$ is such that $u_0^-(O^\delta)\geq u_0^-(O^-)-\delta$. Consequently, we can write $\mathbb{R}$ as a partition

$$\mathbb{R}=F^+\cup(B^-\cap O^\delta)\cup\mathcal{B}^\delta,$$

where $\mathcal{B}^\delta=(F^+\cup(B^-\cap O^\delta))^c$ has small measure $|u_0|(\mathcal{B}^\delta)\leq2\delta$, particles starting in F^+ have a positive sign, and particles starting in $(B^-\cap O^\delta)$ have a negative sign. Let N be large enough to ensure $\varepsilon_N\leq\varepsilon_\delta/3$. The distance between any element of F^+ and any element of O^δ is larger than ε_δ. As a consequence, if the particles with index i and j kill each other before time τ, then either one of them started in $\mathcal{B}^\delta$, or one of the particles i and j moved by a distance larger than $\varepsilon_\delta/3$. This writes

$$\begin{aligned}\mathrm{Card}\left\{i,\kappa_i^N<\tau_\delta\right\}&=2\mathrm{Card}\left\{(i,j),i<j,X^{N,i,r_N}\text{ and }X^{N,j,r_N}\text{ kill each other}\right\}\\&\leq2\mathrm{Card}\left\{i,X_0^{N,i,r_N}\in\mathcal{B}^\delta\text{ or }\sup_{t\in[0,\tau]}|X_t^{N,i,r_N}-X_0^i|\geq\varepsilon_\delta/3\right\}.\end{aligned}$$

As a consequence, if $\tau_\delta>0$ is small enough so that $\mathbb{P}(\sup_{t\in[0,\tau_\delta]}|X_t^{N,i,r_N}-X_0^{N,i,r_N}|\geq\varepsilon_\delta/3)\leq\delta$ (this can be achieved using an adaptation of (7.8)), then

$$\mathbb{P}(\kappa_1^N < \tau_\delta) = \frac{1}{N}\mathbb{E}\left(\text{Card}\left\{i, \kappa_i^N < \tau_\delta\right\}\right) \leq \frac{2}{N}\mathbb{E}\left(\text{Card}\left\{i, X_0^{N,i,r_N} \in \mathcal{B}^\delta \text{ or } \sup_{t\in[0,\tau_\delta]} |X_t^{N,i,r_N} - X_0^i| \geq \varepsilon_\delta/3\right\}\right)$$
$$\leq 2\mathbb{P}\left(X_0^i \in \mathcal{B}^\delta\right) + 2\mathbb{P}\left(\sup_{t\in[0,\tau_\delta]} |X_t^{N,i,r_N} - X_0^{N,i,r_N}| \geq \varepsilon_\delta/3\right)$$
$$\leq 6\delta.$$

Consequently,

$$\mathbb{E}^{\pi^\infty}(Q(\kappa < \tau)) \leq \liminf_N \mathbb{E}^{\pi^N}(Q(\kappa < \tau)) = \liminf_N \mathbb{P}(\kappa_1^N < \tau) \underset{\tau\to 0}{\to} 0.$$

Thus for π^∞-almost all Q, κ is $Q-$almost surely positive. As a consequence, for any bounded continuous function φ,

$$\mathbb{E}^{\pi^\infty}\left|\int_{\mathcal{K}} \mathbf{1}_{\kappa>0}\varphi(f(0))\mathrm{d}Q(\kappa,f) - \int_{\mathbb{R}} \varphi \mathrm{d}|u_0|\right| = \mathbb{E}^{\pi^\infty}\left|\int_{\mathcal{K}} \varphi(f(0))\mathrm{d}Q(\kappa,f) - \int_{\mathbb{R}} \varphi \mathrm{d}|u_0|\right|$$
$$= \lim_N \mathbb{E}^{\pi^N}\left|\int_{\mathcal{K}} \varphi(f(0))\mathrm{d}Q(\kappa,f) - \int_{\mathbb{R}} \varphi \mathrm{d}|u_0|\right| = 0,$$

from the law of large numbers.

The main step in the proof of Theorems 7.3.1, 7.3.2 and 7.3.3 is the following proposition :

Proposition 7.3.6 *Let ε_N and h_N be two sequences going to zero.*

- *If σ_N is a constant sequence and $0 < \alpha \leq 1$, suppose $N^{-1/\alpha} \leq N^{-1/\lambda}\varepsilon_N$ and $N^{-\lambda} \leq 4\sup_{[-1,1]}|A'|h_N \leq \varepsilon_N$ for some positive λ. If $\alpha = 1$, also assume $h_N \leq N^{-1/\lambda}\varepsilon_N$. Then, there exists a sequence (r_N) of positive real numbers, such that the limit of any converging subsequence of π^{N,r_N} gives full measure to the set*

$$\{Q \in \mathcal{P}(\mathcal{K}), H * \tilde{Q}_t(x) \text{ is the entropy solution to (7.1)}\}.$$

- *Let σ_N be a sequence going to zero and assume $N^{-\lambda} \leq 4\sup_{[-1,1]}|A'|h_N \leq \varepsilon_N$ for some positive λ. If $1 < \alpha \leq 2$, also assume $\sigma_N \leq \varepsilon_N^{1-\frac{1}{\alpha}} N^{-1/\lambda}$. Then*

$$\{Q \in \mathcal{P}(\mathcal{K}), H * \tilde{Q}_t(x) \text{ is the entropy solution to (7.2)}\}$$

is given full measure by any limit of a converging subsequence of π^{N,r_N}, for a well chosen sequence (r_N), in the case $\alpha < 2$, and by any limit of a converging subsequence of π^N if $\alpha = 2$.

- *If σ_N is a constant sequence and $1 < \alpha \leq 2$, the limit of any converging subsequence of π^N gives full measure to the set*

$$\{Q \in \mathcal{P}(\mathcal{K}), H * \tilde{Q}_t(x) \text{ is the weak solution to (7.1)}\}.$$

Proposition 7.3.6 will be proved in Section 7.3.1. We first admit it to complete the proofs of Theorems 7.3.1, 7.3.2 and 7.3.3. We need the following lemma.

Lemma 7.3.7 *Let $\alpha < 2$ and r_N be a sequence of positive numbers going to zero. Then, for any $T > 0$,*

$$\lim_{N\to\infty} \int_0^T \mathbb{E}\|H * \tilde{\mu}_t^N - H * \tilde{\mu}_t^{N,r_N}\|_{\mathbb{L}^1\left(\frac{\mathrm{d}x}{1+x^2}\right)} \mathrm{d}t = 0.$$

Proof. By exchangeability of the particles, one has

$$\int_0^T \mathbb{E}\|H*\tilde{\mu}_t^N - H*\tilde{\mu}_t^{N,r_N}\|_{\mathbb{L}^1\left(\frac{\mathrm{d}x}{1+x^2}\right)}\mathrm{d}t \leq \mathbb{E}\int_0^T\int_{\mathbb{R}}\frac{1}{N}\sum_{\kappa_i^N>t}\left|\mathbf{1}_{X_t^{N,i}\leq x}-\mathbf{1}_{X_t^{N,i,r_N}\leq x}\right|\frac{\mathrm{d}x\mathrm{d}t}{x^2+1}$$
$$\leq \int_0^T \mathbb{E}\left(\mathbf{1}_{\kappa_1^N\geq t}\left|X_t^{N,1}-X_t^{N,1,r_N}\right|\wedge\pi\right)\mathrm{d}t.$$

This last quantity goes to zero, since the processes $X^{N,1}$ and $X^{N,1,r_N}$ coincide on the discretization grid, whose mesh goes to zero. Indeed, for $t\in[kh_N,(k+1)h_N)$

$$\mathbb{E}\left(\mathbf{1}_{\kappa_1^N>t}|X_t^{N,1,r_N}-X_t^{N,1}|\wedge\pi\right)\leq \mathbb{E}\left(\mathbf{1}_{\kappa_1^N>t}|X_t^{N,1,r_N}-X_{kh_N}^{N,1}|\wedge\pi\right)+\mathbb{E}\left(\mathbf{1}_{\kappa_1^N>t}|X_t^{N,1}-X_{kh_N}^{N,r_N}|\wedge\pi\right)$$
$$\leq Kh_N^{1/2}. \tag{7.10}$$

For this last estimate, we used, for an $\alpha-$stable Lévy process L, the inequality

$$\mathbb{E}\left(|L_t|\wedge 1\right)\leq K\mathbb{E}\left(|L_t|^{\alpha/2}\right)=Kt^{1/2}.$$

From Lemma 7.3.7, it is sufficient to show $\lim_{N\to\infty}\int_0^T\mathbb{E}\|H*\tilde{\mu}_t^{N,r_N}-v_t\|_{\mathbb{L}^1(\frac{\mathrm{d}x}{1+x^2})}\mathrm{d}t=0$ in order to prove Theorems 7.3.1 and 7.3.2.

Proof (Proof of Theorems 7.3.1-7.3.2-7.3.3). We write the proof for Theorems 7.3.1 and 7.3.2 in the case $\alpha<2$. The proof of Theorem 7.3.2 with $\alpha=2$ and Theorem 7.3.3 is the same, with π^N replacing π^{N,r_N}.

Let γ^k be a Lipschitz continuous approximations of γ, with $\mathbb{P}(\gamma(X_0^1)\neq\gamma^k(X_0^1))\leq 1/k$ (see [36], Lemma 2.5, for a construction of such a γ^k). One has, by exchangeability of the particles,

$$\begin{aligned}&\mathbb{E}\int_0^T\int_{\mathbb{R}}\left|H*\tilde{\mu}_t^{N,r_N}(x)-v_t(x)\right|\frac{\mathrm{d}x\mathrm{d}t}{x^2+1}\\ \leq&\mathbb{E}\int_0^T\int_{\mathbb{R}}\mathbf{1}_{\kappa_1^N>t}H(x-X_t^{N,1,r_N})\left|\gamma(X_0^{N,1,r_N})-\gamma^k(X_0^{N,1,r_N})\right|\frac{\mathrm{d}x\mathrm{d}t}{x^2+1}\\ &+\mathbb{E}^{\pi^N}\left(\int_0^\infty\int_{\mathbb{R}}\left|\int_{\mathcal{K}}\mathbf{1}_{\kappa>t}H(x-f(t))\gamma^k(f(0))\mathrm{d}Q(f,\kappa)-v_t(x)\right|\frac{\mathrm{d}x}{x^2+1}\right).\end{aligned}\tag{7.11}$$

From the assumption on γ^k, the first term in the right hand side of (7.11) is smaller than $2\pi/k$ which tends to zero as k goes to ∞. The bounded function

$$Q\mapsto\int_0^T\int_{\mathbb{R}}\left|\int_{\mathcal{K}}\mathbf{1}_{\kappa>t}H(x-f(t))\gamma^k(f(0))\mathrm{d}Q(f,\kappa)-v_t(x)\right|\frac{\mathrm{d}x\mathrm{d}t}{x^2+1}$$

is continuous. From Proposition 7.3.6, the second term in the right hand side of (7.11) converges, as N goes to ∞ to

$$\mathbb{E}^{\pi^\infty}\left(\int_0^T\int_{\mathbb{R}}\left|\int_{\mathcal{K}}\mathbf{1}_{\kappa>t}H(x-f(t))\left(\gamma^k(f(0))-\gamma(f(0))\right)\mathrm{d}Q(f,\kappa)\right|\frac{\mathrm{d}x}{x^2+1}\right).$$

This terms goes to zero as k tend to infinity using the argument of the begining of the proof with $X^{N,1,r_N}$ replaced by the canonical process y.

7.3.1 Proof of Proposition 7.3.6

This section is devoted to the proof of Proposition 7.3.6. Since the hardest part of this proof is the first two items, we do not give all details for the third item and for the second one in the

case $\alpha = 2$. Indeed, for these two last settings, the separation of small jumps and large jump is not necessary for the proof.

Let r_N be a sequence of positive real numbers, going to zero as $N \to \infty$, which will be explicited later. Let $r > 0$ and c be reals numbers, η a smooth convex function, ψ a primitive of $A'\eta'$ and g a smooth compactly supported nonnegative function. We define the function $\varphi_t(x) = \int_{-\infty}^{x} g_t(y)\mathrm{d}y$. Note that φ is smooth, and nondecreasing with respect to the space variable. We consider a subsequence of π^{N,r_N}, still denoted π^{N,r_N} for simplicity, which converges to a limit π^∞. We want to prove that, for $\pi^\infty-$almost all Q, the function $H*\tilde{Q}_t$ satisfies the entropy formulation associated to the corresponding case.

One can write, for any $k \geq 0$ and $t \in]kh_N, (k+1)h_N]$

$$\mathbb{P}\left(\exists i,j, \kappa_i^N \wedge \kappa_j^N > t, X_t^{N,i,r} = X_t^{N,j,r}\right) = \mathbb{E}\left(\mathbb{P}\left(\exists i,j, \kappa_i^N \wedge \kappa_j^N > t, X_t^{N,i,r} = X_t^{N,j,r} \middle| (X_{kh_N}^{N,q})_q\right)\right)$$
$$= \mathbb{E}\left(\mathbb{P}\left(\exists i,j, \kappa_i^N \wedge \kappa_j^N > t, \sigma_N Z_t^{i,j,k,N} = X_{kh_N}^{N,j} - X_{kh_N}^{N,i} + \mathcal{A}_t^{N,j} - \mathcal{A}_t^{N,i} \middle| (X_{kh_N}^{N,q})_q\right)\right),$$

where we denote

$$Z_t^{i,j,N,k} = \Lambda_t^{N,i,r} - \Lambda_{kh_N}^{N,i,r} - \Lambda_t^{N,j,r} + \Lambda_{kh_N}^{N,j,r} + L_t^{N,i,r} - L_{kh_N}^{N,i,r} - L_t^{N,j,r} + L_{kh_N}^{N,j,r}.$$

From the conditional independence of the processes $L^{N,i,r}$, $L^{N,j,r}$, $\Lambda^{N,i,r}$ and $\Lambda^{N,j,r}$, the random variable $Z_t^{i,j,N,k}$ has a density. As a consequence, since the process $\mathcal{A}_t^{N,j} - \mathcal{A}_t^{N,i}$ is deterministic on $[kh_N, (k+1)h_N]$ conditionally to $(X_{kh_N}^{N,q})_q$, the above probability is zero, meaning that for all time $t > 0$, the alive particles X_t^{N,i,r_N} almost surely have distinct positions. As a consequence, the function $\eta\left(H * \tilde{\mu}_t^{N,r_N}(x)\right)$ is the cumulative distribution function of the signed measure

$$\xi_t^N = \sum_{\kappa_i^N > t} w_t^i \delta_{X_t^{N,i,r_N}},$$

where

$$w_t^i = \mathbf{1}_{\kappa_i^N > t}\left(\eta\left(\frac{1}{N}\sum_{\substack{\kappa_j^N > t \\ X_t^{N,j,r_N} \leq X_t^{N,i,r_N}}} \gamma(X_0^j)\right) - \eta\left(\frac{1}{N}\sum_{\substack{\kappa_j^N > t \\ X_t^{N,j,r_N} < X_t^{N,i,r_N}}} \gamma(X_0^j)\right)\right)$$
$$= \mathbf{1}_{\kappa_i^N > t}\left(\eta\left(H * \tilde{\mu}_t^{N,r_N}\left(X_t^{N,i,r_N}\right)\right) - \eta\left(H * \tilde{\mu}_t^{N,r_N}\left(X_t^{N,i,r_N}-\right)\right)\right).$$

Let $(\zeta_m)_{m\in\mathbb{N}}$ be the increasing sequence of times which are either a jump time for some L^{N,i,r_N} (*i.e.* a jump of size $> r_N$ for X^{N,i,r_N}) or either a time of the form $kh_N/2$. One has

$$-\left\langle \xi_0^N, \varphi_0 \right\rangle = \sum_{m=1}^{\infty} \left\langle \xi_{\zeta_m}^N, \varphi_{\zeta_m} \right\rangle - \left\langle \xi_{\zeta_{m-1}}^N, \varphi_{\zeta_{m-1}} \right\rangle$$
$$= \sum_{\kappa_i^N > 0} \sum_{m=1}^{\infty} w_{\zeta_{m-1}}^i \left(\varphi_{\zeta_m}\left(X_{\zeta_m-}^{N,i,r_N}\right) - \varphi_{\zeta_{m-1}}\left(X_{\zeta_{m-1}}^{N,i,r_N}\right)\right) \tag{7.12}$$
$$+ \sum_{\kappa_i^N > 0} \sum_{m=1}^{\infty} \left(w_{\zeta_m}^i \varphi_{\zeta_m}\left(X_{\zeta_m}^{N,i,r_N}\right) - w_{\zeta_{m-1}}^i \varphi_{\zeta_m}\left(X_{\zeta_m-}^{N,i,r_N}\right)\right).$$

Notice that these infinite sums are actually finite, since the function φ_t is identically zero when t is large enough, and since the process $(L^{N,1,r_N},\dots,L^{N,N,r_N})$ has a finite number of jumps on bounded intervals.

We consider the first term in the right hand side of (7.12). Denote by $\nu^{i,r}=\sum_{\Delta X_t^{N,i,r}\neq 0}\delta_{(\Delta L_t^{N,i,r}+\Delta A_t^{N,i,r},t)}$ the jump measure associated to $L^{N,i,r}+A^{N,i,r}$, and by

$$\tilde\nu^{i,r}(\mathrm{d}y,\mathrm{d}t)=\nu^{i,r}(\mathrm{d}y,\mathrm{d}t)-2c_\alpha\left(\chi_t^N\mathbf{1}_{|y|\le r}+(1-\chi_t^N)\mathbf{1}_{|y|>r}\right)\frac{\mathrm{d}y\mathrm{d}t}{|y|^{1+\alpha}}$$

its compensated measure, where $\chi_t^N=\sum_{k=0}^\infty \mathbf{1}_{[kh_N,(k+1/2)h_N)}(t)$. Let us apply Itō's Formula on the interval (ζ_{m-1},ζ_m). If $\zeta_{m-1}=kh_N$ for some integer k, then $\zeta_m=(k+1/2)h_N$, and almost surely $X^{N,i,r}_{(k+\frac12)h_N-}=X^{N,i,r}_{(k+\frac12)h_N}$ holds. As a consequence

$$\begin{aligned}
&\varphi_{(k+\frac12)h_N}\left(X^{N,i,r}_{(k+\frac12)h_N-}\right)-\varphi_{kh_N}\left(X^{N,i,r}_{kh_N}\right)\\
&=\int_{kh_N}^{(k+\frac12)h_N}\partial_t\varphi_t(X_t^{N,i,r})\mathrm{d}t+2\int_{kh_N}^{(k+\frac12)h_N}\partial_x\varphi_t(X_t^{N,i,r})A'\left(H*\tilde\mu^{N,r}_{kh_N}(X^{N,i,r}_{kh_N})\right)\mathrm{d}t\\
&\quad+\int_{(kh_N,(k+1/2)h_N)}\int_{\{|y|\le r\}}\left(\varphi_t(X^{N,i,r}_{t-}+\sigma_N y)-\varphi_t(X^{N,i,r}_{t-})-\sigma_N y\partial_x\varphi_t(X^{N,i,r}_{t-})\right)\nu^{i,r}(\mathrm{d}y,\mathrm{d}t)\\
&\quad+\sigma_N\int_{(kh_N,(k+1/2)h_N)}\partial_x\varphi_t(X^{N,i,r}_{t-})\left(\int_{\{|y|\le r\}}y\tilde\nu^{i,r}(\mathrm{d}y,\mathrm{d}t)\right).
\end{aligned}$$

If ζ_{m-1} is not of the form kh_N, then the process $X^{N,i,r}$ is constant on the interval $[\zeta_{m-1},\zeta_m)$, and one has $\varphi_{\zeta_m}(X^{N,i,r}_{\zeta_m-})-\varphi_{\zeta_{m-1}}(X^{N,i,r}_{\zeta_{m-1}})=\int_{\zeta_{m-1}}^{\zeta_m}\partial_t\varphi_t(X_t^{N,i,r})\mathrm{d}t$. Summing over all the intervals (ζ_{m-1},ζ_m), Equation (7.12) writes, denoting $\tau_t=\max\{\zeta_m,\zeta_m\le t\}$,

$$\begin{aligned}
-\left\langle\xi_0^N,\varphi_0\right\rangle&=\sum_{\kappa_i^N>0}\int_0^\infty w^i_{\tau_t}\left(\partial_t\varphi_t(X_t^{N,i,r_N})+2\chi_t^N\partial_x\varphi_t(X_t^{N,i,r_N})A'\left(H*\tilde\mu^{N,r_N}_{\tau_t}(X^{N,i,r_N}_{\tau_t})\right)\right)\mathrm{d}t\\
&\quad+c_\alpha\sum_{\kappa_i^N>0}\int_0^\infty w^i_{\tau_t}\chi_t^N\int_{\{|y|\le r_N\}}\left(\varphi_t(X_t^{N,i,r_N}+\sigma_N y)-\varphi_t(X_t^{N,i,r_N})-\sigma_N y\partial_x\varphi_t(X_t^{N,i,r_N})\right)\frac{2\mathrm{d}y\mathrm{d}t}{|y|^{1+\alpha}}\\
&\quad+\sum_{\kappa_i^N>0}\sum_{\substack{\text{large jump}\\ \text{at }\zeta_m}}w^i_{\zeta_m}\varphi_{\zeta_m}(X^{N,i,r_N}_{\zeta_m})-w^i_{\zeta_{m-1}}\varphi_{\zeta_m}(X^{N,i,r_N}_{\zeta_m-})\\
&\quad+\sum_{\kappa_i^N>0}\sum_{\substack{\zeta_m\text{ of the}\\ \text{form }kh_N}}(w^i_{\zeta_m}-w^i_{\zeta_{m-1}})\varphi_{\zeta_m}(X^{N,i,r_N}_{\zeta_m})\qquad(7.13)\\
&\quad+\sum_{\kappa_i^N>0}\sum_{\substack{\zeta_m\text{ of the}\\ \text{form }(k+1/2)h_N}}(w^i_{\zeta_m}-w^i_{\zeta_{m-1}})\varphi_{\zeta_m}(X^{N,i,r_N}_{\zeta_m})\\
&\quad+M_N.
\end{aligned}$$

Here, the third, fourth and fifth terms correspond to the second term in the right hand side of (7.12), and M_N is a martingale term given by

$$M_N=\sum_{\kappa_i^N>0}\int_0^\infty w^i_{\tau_t}\chi_t^N\int_{\{|y|\le r_N\}}\left(\varphi_t(X^{N,i,r_N}_{t-}+\sigma_N y)-\varphi_t(X^{N,i,r_N}_{t-})\right)\tilde\nu^{i,r_N}(\mathrm{d}y,\mathrm{d}t).$$

Equation (7.13) can be rewritten

$$T_N^1 = T_N^2 + T_N^3 + T_N^4 + T_N^5 + M_N,$$

where $T_N^1 = -\langle \xi_0^N, \varphi_0 \rangle$, T_N^2 is the sum of the two first terms in the right-hand-side of (7.13), T_N^3 is the third one, T_N^4 the fourth one and T_N^5 the fifth one.

The four following Lemmas, whose proofs are postponed to Section 7.3.2 deal with the asymptotic behavior of the terms M_N, $T_N^2 - T_N^1$, T_N^3 and T_N^4.

Lemma 7.3.8 *For some positive constant K,*

$$\mathbb{E}|M_N|^2 \leq \frac{K\sigma_N^2 r_N^{2-\alpha}}{N}$$

holds. The equivalent term in the case $\alpha = 2$,

$$M_N = \sigma_N \sum_{\kappa_i^N > 0} \int_0^\infty w_{\tau_t}^i \partial_x \varphi(X_t^{N,i}) \mathrm{d}L_t^i,$$

satifies the same estimate :

$$\mathbb{E}|M_N|^2 \leq K\frac{\sigma_N^2}{N}.$$

Lemma 7.3.9 – *It holds that*

$$\mathbb{E}\Bigg| - T_N^1 + T_N^2 + \int_0^\infty \int_{\mathbb{R}} \Big(\eta(H * \tilde{\mu}_t^{N,r_N})\partial_t g_t + 2\chi_t^N \psi(H * \tilde{\mu}_t^{N,r_N})\partial_x g_t\Big)\,\mathrm{d}t + \int_{\mathbb{R}} g_0 \eta(H * \tilde{\mu}_0^{N,r_N})\mathrm{d}x$$
$$+2c_\alpha \int_0^\infty \chi_t^N \int_{\mathbb{R}} \int_{\{|y|\leq r_N\}} \eta(H * \tilde{\mu}_t^{N,r_N}(x))\left(g_t(x+\sigma_N y) - g_t(x) - \sigma_N y \partial_x g_t(x)\right)\frac{\mathrm{d}y\mathrm{d}x\mathrm{d}t}{|y|^{1+\alpha}}\Bigg| \underset{N\to\infty}{\to} 0.$$

– *If $r_N \leq 1/\sigma_N$, then*

$$\left|2c_\alpha \int_0^\infty \chi_t^N \int_{\mathbb{R}} \int_{\{|y|\leq r_N\}} \eta(H * \tilde{\mu}_t^{N,r_N}(x))\left(g_t(x+\sigma_N y) - g_t(x) - \sigma_N y \partial_x g_t(x)\right)\frac{\mathrm{d}y\mathrm{d}x\mathrm{d}t}{|y|^{1+\alpha}}\right| \leq K\sigma_N^\alpha.$$

The following lemma gives two estimates for the term T_N^3, the first being useful for a constant viscosity $\sigma_N \equiv \sigma$, and the second for vanishing viscosity $\sigma_N \to 0$.

Lemma 7.3.10 – *The error term*

$$\mathbb{E}\left|T_N^3 + 2c_\alpha \int_0^\infty (1-\chi_t^N) \int_{\mathbb{R}} \int_{\{|y|>r_N\}} \eta'(H * \tilde{\mu}_t^{N,r_N}(x))\left(H * \tilde{\mu}_t^{N,r_N}(x+\sigma_N y) - H * \tilde{\mu}_t^{N,r_N}(x)\right) g_t(x)\frac{\mathrm{d}y\mathrm{d}x\mathrm{d}t}{|y|^{1+\alpha}}\right|$$

goes to zero if $N^{-1}r_N^{-\alpha}$ goes to 0.

– *It holds that*

$$\mathbb{E}|T_N^3| \leq K(\sigma_N r_N^{1-\alpha} + \sigma_N^\alpha).$$

Lemma 7.3.11 *One has $\mathbb{E}|T_N^4| \underset{N\to\infty}{\to} 0$.*

We now have to control the probability for the last remaining term T_N^5 to be negative. If there is no crossing of particles with opposite signs between kh_N and $(k+1/2)h_N$, for any k, then $T_N^5 \geq 0$. Indeed, let $X_{(k+1/2)h_N}^{N,i_1,r_N} \leq \ldots \leq X_{(k+1/2)h_N}^{N,i_q,r_N}$ be a maximal sequence of consecutive particles with same sign. The sequence $\left(\varphi_{(k+1/2)h_N}(X_{(k+1/2)h_N}^{N,i_l,r_N})\right)_{l=1,\ldots,q}$ is thus a nondecreasing sequence, and from the convexity of η and the fact that no particles with opposite signs cross, $(w_{(k+1/2)h_N}^{i_l})_{l=1,\ldots,q}$ is the nondecreasing reordering of $(w_{kh_N}^{i_l})_{l=1,\ldots,q}$. Thus, from Lemma 7.3.13 below, $\sum_{\kappa_i^N > kh_N}(w_{(k+1/2)h_N}^i - w_{kh_N}^i)\varphi_{(k+1/2)h_N}(X_{(k+1/2)h_N}^{N,i,r_N})$ is nonnegative. It is thus sufficient to

control the probability that two particles with opposite signs cross between kh_N and $(k+1/2)h_N$. Since after the murder there is no couple of particles with opposite signs separated by a smaller distance than ε_N, this does not happen as soon as no particle drift by more than $\varepsilon_N/4$ and no particle is moved by more than $\varepsilon_N/4$ by the small jumps. The drift on half a time step is smaller than $\sup_{[-1,1]}|A'|h_N$ which is assumed to be smaller than $\varepsilon_N/4$. We control the contribution of the small jumps in the following lemma :

Lemma 7.3.12 *Let B_N be the event*

$$B_N = \left\{\forall k \leq T/h_N, \forall i, \sigma_N \left|\Lambda^{i,r_N}_{(k+1/2)h_N} - \Lambda^{i,r_N}_{kh_N}\right| \leq \varepsilon_N/4\right\},$$

so that no crossing of particles with opposite signs between kh_N and $(k+1/2)h_N$ occurs on B_N. One has, for $\alpha < 2$,

$$\mathbb{P}(B_N) \geq \left(1 - exp\left(Kh_N r_N^{-\alpha} - \varepsilon_N/4\sigma_N r_N\right)\right)^{NT/h_N},$$

For $\alpha = 2$, we define the event B_N by

$$B_N = \left\{\forall k \leq T/h_N, \forall i, \sigma_N \left|L^i_{(k+1)h_N} - L^i_{kh_N}\right| \leq \varepsilon_N/4\right\}.$$

Then, one has

$$\mathbb{P}(B_N) \geq \left(1 - K\,exp\left(-\varepsilon_N^2/(32h_N\sigma_N^2)\right)\right)^{NT/h_N}.$$

The proof will be given in Section 7.3.2.

We now gather all the previous information to prove that, depending on the considered case, the entropic formulation or the weak formulation holds almost surely.

1. *Constant viscosity $\sigma_N \equiv \sigma$, with index $0 < \alpha \leq 1$.*
   Define, for $Q \in \mathcal{P}(\mathcal{K})$,

$$\begin{aligned} F^r_N(Q) = &\int_{\mathbb{R}} \eta(H*\tilde{Q}_0)g_0 + \int_0^\infty \int_{\mathbb{R}} \left(\eta(H*\tilde{Q}_t)\partial_t g + 2\chi^N_t \psi(H*\tilde{Q}_t)\partial_x g\right) \mathrm{d}t \\ &+ 2c_\alpha \int_0^\infty (1-\chi^N_t) \int_{\mathbb{R}} \int_{\{|y|>r\}} \eta'(H*\tilde{Q}_t(x))(H*\tilde{Q}_t(x+\sigma_N y) - H*\tilde{Q}_t(x))g_t(x) \frac{\mathrm{d}y\mathrm{d}x\mathrm{d}t}{|y|^{1+\alpha}} \\ &+ 2c_\alpha \int_0^\infty \chi^N_t \int_{\mathbb{R}} \int_{\{|y|\leq r\}} \eta(H*\tilde{Q}_t(x))(g_t(x+\sigma_N y) - g_t(x) - \sigma_N y \partial_x g_t(x)) \frac{\mathrm{d}y\mathrm{d}x\mathrm{d}t}{|y|^{1+\alpha}} \end{aligned}$$

   and

$$\begin{aligned} F^r(Q) = &\int_{\mathbb{R}} \eta(H*\tilde{Q}_0)g_0 + \int_0^\infty \int_{\mathbb{R}} \left(\eta(H*\tilde{Q}_t)\partial_t g + \psi(H*\tilde{Q}_t)\partial_x g\right) \mathrm{d}t \\ &+ c_\alpha \int_0^\infty \int_{\mathbb{R}} \int_{\{|y|>r\}} \eta'(H*\tilde{Q}_t(x))(H*\tilde{Q}_t(x+\sigma y) - H*\tilde{Q}_t(x))g_t(x) \frac{\mathrm{d}y\mathrm{d}x\mathrm{d}t}{|y|^{1+\alpha}} \\ &+ c_\alpha \int_0^\infty \int_{\mathbb{R}} \int_{\{|y|\leq r\}} \eta(H*\tilde{Q}_t(x))(g_t(x+\sigma y) - g_t(x) - \sigma y \partial_x g_t(x)) \frac{\mathrm{d}y\mathrm{d}x\mathrm{d}t}{|y|^{1+\alpha}}. \end{aligned}$$

   Notice that from the convexity of η, one has

$$\eta'(H*\tilde{Q}_t(x))(H*\tilde{Q}_t(x+\sigma y) - H*\tilde{Q}_t(x)) \leq \eta(H*\tilde{Q}_t(x+\sigma y)) - \eta(H*\tilde{Q}_t(x)),$$

   so that for any $0 < r \leq r'$, one has $F^r \leq F^{r'}$ and $F^r_N \leq F^{r'}_N$.
   From Equation (7.13), it holds, for N large enough so that $r_N \leq r$, that

$$F^r_N(\mu^{N,r_N}) \geq F^{r_N}_N(\mu^{N,r_N}) = T^5_N + \left(-T^1_N + T^2_N + T^3_N + T^4_N + M_N + F^{r_N}_N(\mu^{N,r_N})\right).$$

From the assumptions made on ε_N and h_N one can construct a sequence r_N such that $N^{-1/\alpha} = o(r_N)$, $h_N r_N^{-\alpha} = o\left(\varepsilon_N r_N^{-1}\right)$ and $\frac{N}{h_N}\exp\left(-\varepsilon_N/4\sigma r_N\right) \to 0$. Indeed, set $r_N = \varepsilon_N N^{-1/2\lambda}$. Then, one has $N^{-1/\alpha} r_N^{-1} \leq K N^{-1/2\lambda}$ and $\frac{h_N}{\varepsilon_N} r_N^{1-\alpha} = h_N \varepsilon_N^{-\alpha} N^{(\alpha-1)/2\lambda}$, this last term going to zero for any value of α. Then $\frac{N}{h_N}$ goes to infinity at the rate of a power of N, and $\varepsilon_N / r_N = N^{1/2\lambda}$ as well. Thus, $\frac{N}{h_N}\exp\left(-\varepsilon_N/4\sigma r_N\right)$ tends to zero.

As a consequence, from Lemmas 7.3.8, 7.3.9, 7.3.10 and 7.3.11, $\mathbb{E}\left|-T_N^1 + T_N^2 + T_N^3 + T_N^4 + M_N + F_N^{r_N}(\mu^{N,r_N})\right|$ goes to zero as N tends to infinity, and the event B_N defined in Lemma 7.3.12 is such that $\mathbb{P}(B_N) \to 1$. On the event B_N, T_N^5 is almost-surely nonnegative, so that, from the uniform boundedness of F_N^r with respect to N, $\mathbb{E}^{\pi^{N,r_N}}(F_N^r(Q)^-) = \mathbb{E}(F_N^r(\mu^{N,r_N})^-)$ goes to 0. To show that the entropic formulation holds almost surely, we need a continuous approximation of F_N^r and F^r. We define $F^{r,\delta}$ and $F_N^{r,\delta}$ by replacing every occurrence of $H * \tilde{Q}_t$ in the definitions of F^r and F_N^r by $\int_{\mathcal{K}} \mathbf{1}_{\kappa>t} H(.-f(t))\gamma^\delta(f(0))\mathrm{d}Q(f,\kappa)$, where γ^δ is a Lipschitz continuous approximation of γ, with $\mathbb{P}(\gamma(X_0^1) \neq \gamma^\delta(X_0^1)) \leq \delta$ (see [36], Lemma 2.5, for the construction of γ^δ). Then, for any fixed δ and r, the family $\{F^{r,\delta}\} \cup \{F_N^{r,\delta}, N \in \mathbb{N}\}$ is equicontinuous for the topology of weak convergence. Indeed, let Q^k be a sequence of probability measures on $\mathcal{K}$ converging to Q as k goes to infinity. From the continuity of the application $\mathcal{K} \to \mathbb{R}, (f,\kappa) \mapsto \mathbf{1}_{\kappa>0} f(0)$, Q_0^k converges weakly to Q_0 (where Q_0 and Q_0^k are defined as in (7.9)), and from the continuity of the applications $\mathcal{K} \to \mathbb{R}, (f,\kappa) \mapsto \mathbf{1}_{\kappa>t}\gamma^\delta(f(0))\mathbf{1}_{f(t)\leq y}$ on the set $\{(f,\kappa) \in \mathcal{K}, f(t) = f(t-), f(t) \neq y\}$, for all t in the complement of the countable set $\{t \in [0,\infty), Q(\{f(t) \neq f(t-)\} \cup \{\kappa = t\}) > 0\}$, the quantity $\int_{\mathcal{K}} \mathbf{1}_{\kappa>t} H(.-f(t))\gamma^\delta(f(0))\mathrm{d}Q^k(f,\kappa)$ converges almost everywhere to $\int_{\mathcal{K}} \mathbf{1}_{\kappa>t} H(.-f(t))\gamma^\delta(f(0))\mathrm{d}Q(f,\kappa)$. From Lebesgue's bounded convergence theorem, we deduce that

$$\sup_N |F_N^{r,\delta}(Q^k) - F_N^{r,\delta}(Q)| + |F^{r,\delta}(Q^k) - F^{r,\delta}(Q)| \underset{k\to\infty}{\to} 0$$

yielding equicontinuity for $\{F^{r,\delta}\} \cup \{F_N^{r,\delta}, N \in \mathbb{N}\}$. Moreover, since the sequence χ_t^N converges $*$-weakly to $1/2$ in the space $\mathbb{L}^\infty((0,\infty))$, $F_N^{r,\delta}$ converges pointwise to $F^{r,\delta}$ as N goes to infinity. Ascoli's theorem thus implies that $F_N^{r,\delta}$ converges uniformly on compact sets to $F^{r,\delta}$. From the weak convergence of π^{N,r_N} to π^∞, one thus deduces

$$\mathbb{E}^{\pi^{N,r_N}}[F_N^{r,\delta}(Q)^-] \underset{N\to\infty}{\to} \mathbb{E}^{\pi^\infty}[F^{r,\delta}(Q)^-].$$

Moreover, for any $t > 0$, any y, and any probability measure Q satisfying $Q_0 = |u_0|$ (with Q_0 defined as in (7.9)), which holds true for $\pi^\infty-$almost all Q from Lemma 7.3.5, one has

$$\left|H * \tilde{Q}_t(y) - \int_{\mathcal{K}} \mathbf{1}_{\kappa>t} H(y-f(t))\gamma^\delta(f(0))\mathrm{d}Q(f,\kappa)\right| \leq \int_{\mathbb{R}} |\gamma - \gamma^\delta|\mathrm{d}|u_0| \leq \delta,$$

yielding convergence to 0 for $\mathbb{E}^{\pi^\infty}|F^r(Q)^- - F^{r,\delta}(Q)^-| + \mathbb{E}^{\pi^{N,r_N}}|F_N^r(Q)^- - F_N^{r,\delta}(Q)^-|$ as δ goes to 0, uniformly in N. As a consequence, writing

$$\begin{aligned}\mathbb{E}^{\pi^\infty}(F^r(Q)^-) \leq & \mathbb{E}^{\pi^\infty}|F^r(Q)^- - F^{r,\delta}(Q)^-| + \left|\mathbb{E}^{\pi^\infty}(F^{r,\delta}(Q)^-) - \mathbb{E}^{\pi^{N,r_N}}(F_N^{r,\delta}(Q)^-)\right| \\ & + \mathbb{E}^{\pi^{N,r_N}}|F_N^{r,\delta}(Q)^- - F_N^r(Q)^-| + \mathbb{E}^{\pi^{N,r_N}}(F_N^r(Q)^-)\end{aligned}$$

we deduce that $F^r(Q)$ is nonnegative for $\pi^\infty-$almost all Q. We just have to notice that Lemma 7.3.5 yields that, $\pi^\infty-$almost surely, $H * \tilde{Q}_0 = v_0$ to conclude that the entropy formulation holds $\pi^\infty-$almost surely.

2. *Vanishing viscosity* $\sigma_N \to 0$.
We define

$$F_N^r(Q) = \int_{\mathbb{R}} \eta(H*\tilde{Q}_0)g_0 + \int_0^\infty \int_{\mathbb{R}} \left(\eta(H*\tilde{Q}_t)\partial_t g + 2\chi_t^N \psi(H*\tilde{Q}_t)\partial_x g\right) \mathrm{d}t$$

and

$$F(Q) = \int_{\mathbb{R}} \eta(H*\tilde{Q}_0)g_0 + \int_0^\infty \int_{\mathbb{R}} \left(\eta(H*\tilde{Q}_t)\partial_t g + \psi(H*\tilde{Q}_t)\partial_x g\right) \mathrm{d}t.$$

Regularized versions $F_N^{r,\delta}$ and F^δ of F_N^r and F are also considered using the function γ^δ instead of γ. In the case $\alpha < 2$, the same arguments as above, using the second parts of Lemmas 7.3.9 and 7.3.10 will show that the entropy formulation holds $\pi^\infty-$almost surely for $H*\tilde{Q}_t$, provided there exists a sequence r_N such that $\frac{\sigma_N^2 r_N^{2-\alpha}}{N}$ and $\sigma_N r_N^{1-\alpha}$ go to zero, $r_N \leq \sigma_N^{-1}$, $h_N r_N^{-\alpha} = o(\varepsilon_N(\sigma_N r_N)^{-1})$ and $\frac{N}{h_N}\exp\left(-\varepsilon_N/4\sigma_N r_N\right) \to 0$.

- For $\alpha \leq 1$, any sequence r_N going to zero with a very quick rate will fit.
- For $\alpha > 1$, since we assumed $\sigma_N \leq \varepsilon_N^{1-\frac{1}{\alpha}} N^{-1/\lambda}$ these conditions are satisfied by the sequence $r_N = \frac{\varepsilon_N}{\sigma_N} N^{-\frac{\alpha}{2\lambda(\alpha-1)}}$.

In the case $\alpha = 2$, Itō's formula writes

$$\begin{aligned}
\varphi_{(k+1)h_N}\left(X^{N,i}_{(k+1)h_N}\right) - \varphi_{kh_N}\left(X^{N,i}_{kh_N}\right) = &\int_{kh_N}^{(k+1)h_N} \partial_t\varphi_t(X_t^{N,i})\mathrm{d}t \\
&+ 2\int_{kh_N}^{(k+1)h_N} \partial_x\varphi_t(X_t^{N,i})A'\left(H*\tilde{\mu}^N_{kh_N}(X^{N,i}_{kh_N})\right)\mathrm{d}t. \\
&+ \sigma_N^2 \int_{(kh_N,(k+1)h_N)} \int_{\{|y|\leq r\}} \partial_x^2\varphi_t(X_t^{N,i})\mathrm{d}t \\
&+ \sigma_N \int_{(kh_N,(k+1)h_N)} \partial_x\varphi_t(X_t^{N,i})\mathrm{d}L_t^i.
\end{aligned}$$

The three first terms are treated as in the case $\alpha < 2$, and the stochastic integral is dealt with using Lemma 7.3.8. For the entropic inequality to holds, we need to control the crossing of particles with opposite sign. From Lemma 7.3.12, if $\frac{N}{h_N}\exp\left(-\frac{\varepsilon_N^2}{32\sigma_N^2 h_N}\right)$ goes to zero, then no crossing occurs. Since our assumptions yield $h_N\sigma_N^2 \leq \varepsilon_N^2 N^{-1/\lambda}$ and $N/h_N \leq KN^{1+\lambda}$, this condition holds true.

3. *Constant viscosity $\sigma_N \equiv \sigma$, with index $1 < \alpha \leq 2$.*

In this case, since we want to derive a weak formulation, we do not need to consider separately large and small jumps. As a consequence it is enough to study the process $X_t^{N,i}$.

Let g be a smooth function with compact support, and define for $Q \in \mathcal{P}(\mathcal{K})$,

$$F(Q) = \int_{\mathbb{R}} H*\tilde{Q}_0 g_0 + \int_0^\infty \int_{\mathbb{R}} H*\tilde{Q}_t\partial_t g_t \mathrm{d}t - \sigma^\alpha \int_0^\infty \int_{\mathbb{R}} H*\tilde{Q}_t(-\Delta)^{\frac{\alpha}{2}} g_t \mathrm{d}t + \int_0^\infty \int_{\mathbb{R}} A(H*\tilde{Q}_t)\partial_x g_t.$$

Let $\varphi_t(x) = \int_{-\infty}^x g_t(y)\mathrm{d}y$. One has

$$\begin{aligned}
-\frac{1}{N}\sum_{\kappa_i^N>0} \gamma(X_0^{N,i})\varphi_0(X_0^{N,i}) = &-\frac{1}{N}\sum_{k=0}^\infty \sum_{\kappa_i^N=(k+1)h_N} \gamma(X_0^{N,i})\varphi_{(k+1)h_N}(X^{N,i}_{(k+1)h_N}) \\
&+ \frac{1}{N}\sum_{k=0}^\infty \sum_{\kappa_i^N>kh_N} \gamma(X_0^{N,i})\left(\varphi_{(k+1)h_N}(X^{N,i}_{(k+1)h_N}) - \varphi_{kh_N}(X^{N,i}_{kh_N})\right).
\end{aligned}$$

From Itō's formula, in the case $\alpha < 2$, when $\kappa_i^N > kh_N$,

$$
\begin{aligned}
&\varphi_{(k+1)h_N}(X^{N,i}_{(k+1)h_N}) - \varphi_{kh_N}(X^{N,i}_{kh_N}) \qquad (7.14)\\
=&\int_{kh_N}^{(k+1)h_N} \partial_t\varphi_t(X^{N,i}_t)\mathrm{d}t + \int_{kh_N}^{(k+1)h_N} \partial_x\varphi_t(X^{N,i}_t)A'\left(H*\tilde{\mu}_{kh_N}(X^{N,i}_{kh_N})\right)\mathrm{d}t\\
&+ c_\alpha \int_{(kh_N,(k+1)h_N)}\int_{\mathbb{R}} \left(\varphi_t(X^{N,i}_t+\sigma y) - \varphi_t(X^{N,i}_t) - \mathbf{1}_{\{|y|\le r\}}\sigma y\partial_x\varphi_t(X^{N,i}_t)\right)\frac{\mathrm{d}y\mathrm{d}t}{|y|^{1+\alpha}}\\
&+ \int_{(kh_N,(k+1)h_N)}\int_{\mathbb{R}} \left(\varphi_t(X^{N,i}_{t-}+\sigma y) - \varphi_t(X^{N,i}_{t-})\right)\tilde{\nu}^i(\mathrm{d}y,\mathrm{d}t).
\end{aligned}
$$

We denote $\tau_t = \max\{kh_N, kh_N \le t\}$. Multiplying (7.14) by $\frac{1}{N}\mathbf{1}_{\kappa_i^N > kh_N}\gamma(X_0^{N,i})$, summing over i and k, and integrating by parts, one obtains

$$
\begin{aligned}
\int_{\mathbb{R}} g_0 H*\tilde{\mu}_0^N =& -\int_0^\infty\int_{\mathbb{R}} \partial_t g_t H*\tilde{\mu}_t^N\mathrm{d}t + \int_0^\infty\int_{\mathbb{R}} (-\Delta)^{\frac{\alpha}{2}} g_t H*\tilde{\mu}_t^N\mathrm{d}t\\
&+\frac{1}{N}\int_0^\infty \sum_{\kappa_i^N>\tau_t} \gamma(X_0^{N,i})\partial_x\varphi_t(X_t^{N,i})A'\left(H*\tilde{\mu}_{kh_N}(X^{N,i}_{kh_N})\right)\mathrm{d}t\\
&+\frac{1}{N}\int_{(0,\infty)\times\mathbb{R}} \sum_{\kappa_i^N>\tau_t} \gamma(X_0^{N,i})\left(\varphi_t(X^{N,i}_{t-}+\sigma y) - \varphi_t(X^{N,i}_{t-})\right)\tilde{\nu}^i(\mathrm{d}y,\mathrm{d}t)\\
&-\frac{1}{N}\sum_{k=0}^\infty \sum_{\kappa_i^N=(k+1)h_N} \gamma(X_0^{N,i}))\varphi_{(k+1)h_N}(X^{N,i}_{(k+1)h_N}), \qquad (7.15)
\end{aligned}
$$

Combining an adaptation of Lemma 7.3.14, stated in Section 7.3.2, with A replacing η, and integrating by parts, the difference

$$
\frac{1}{N}\int_0^\infty \sum_{\kappa_i^N>\tau_t} \gamma(X_0^{N,i})\partial_x\varphi_t(X_t^{N,i})A'\left(H*\tilde{\mu}_{kh_N}(X^{N,i}_{kh_N})\right)\mathrm{d}t + \int_0^\infty\int_{\mathbb{R}} \partial_x g_t A(H*\tilde{\mu}_t^N)\mathrm{d}t
$$

tends to zero in $\mathbb{L}^1$. Using an adaptation Lemma 7.3.8, the the fourth term in the right hand side of (7.15) goes to zero in $\mathbb{L}^2$. The fifth term tends to zero in $\mathbb{L}^1$ since

$$
\begin{aligned}
&\left|\frac{1}{N}\sum_{k=0}^\infty \sum_{\kappa_i^N=(k+1)h_N} \gamma(X_0^{N,i})\varphi_{(k+1)h_N}(X^{N,i}_{(k+1)h_N})\right|\\
\le&\frac{1}{N}\sum_{k=0}^\infty \sum_{\substack{\text{pairs } \{i,j\} \text{ killled}\\ \text{at time } (k+1)h_N}} \left|\varphi_{(k+1)h_N}\left(X^{N,i}_{(k+1)h_N}\right) - \varphi_{(k+1)h_N}\left(X^{N,j}_{(k+1)h_N}\right)\right|\\
\le& K\varepsilon_N.
\end{aligned}
$$

As a consequence, $\mathbb{E}^{\pi^N}|F(Q)| = \mathbb{E}|F(\mu^N)|$ tends to zero. We conclude by regularizing the function γ as in the two first points, that $\mathbb{E}^{\pi^\infty}|F(Q)| = 0$. Thus, $F(Q) = 0$ almost surely, so that $H*\tilde{Q}$ almost surely satisfies the weak formulation.
The case $\alpha = 2$ is treated in the same way, the only difference lying in the nature of the stochastic integral.

7.3.2 Proofs of Lemmas 7.3.8 to 7.3.12

In this section, we give the proofs of the previously admitted lemmas of Section 7.3.1.

Proof (Proof of Lemma 7.3.8). Since the particles are driven by independent stable processes and since the inequality $|w_t^i| \le \frac{K}{N}$ holds for some constant K not depending on t, i and N,

$$\begin{aligned}\mathbb{E}M_N^2 =&\mathbb{E}\left|\sum_{\kappa_i^N>0}\int_0^\infty w^i_{\tau_t}\chi_t^N\left(\int_{\{|y|\le r_N\}}\left(\varphi_t(X^{N,i,r_N}_{t-}+\sigma_N y)-\varphi_t(X^{N,i,r_N}_{t-})\right)\tilde\nu^{i,r_N}(\mathrm{d}y,\mathrm{d}t)\right)\right|^2\\ \le&2\sigma_N^2 c_\alpha\mathbb{E}\left(\sum_{\kappa_i^N>0}\int_0^\infty (w^i_{\tau_t})^2\chi_t^N\int_{\{|y|\le r_N\}}(y\|g_t\|_\infty)^2\frac{\mathrm{d}y\mathrm{d}t}{|y|^{1+\alpha}}\right)\\ \le&K\frac{\sigma_N^2 r_N^{2-\alpha}}{N}\int_0^\infty\|g_t\|_\infty^2\mathrm{d}t.\end{aligned}$$

A similar proof with stochastic integrals against Brownian motion yields the result for $\alpha=2$.

Proof (Proof of Lemma 7.3.9). Integrating by parts, one finds

$$\sum_{i=1}^N\int_0^\infty w_t^i\partial_t\varphi_t\left(X_t^{N,i,r_N}\right)\mathrm{d}t=-\int_0^\infty\int_{\mathbb{R}}\eta(H*\tilde\mu_t^{N,r_N})\partial_t g_t\mathrm{d}t+\int_0^\infty\int_{\mathbb{R}}\eta(\tilde\mu_t^{N,r_N}(\mathbb{R}))\partial_t g_t\mathrm{d}t$$

yielding, from Lemma 7.3.14 below,

$$\mathbb{E}\left|\sum_{i=1}^N\int_0^\infty w^i_{\tau_t}\partial_t\varphi_t\left(X_t^{N,i,r_N}\right)\mathrm{d}t+\int_0^\infty\int_{\mathbb{R}}\eta(H*\tilde\mu_t^{N,r_N})\partial_t g_t\mathrm{d}t-\int_0^\infty\int_{\mathbb{R}}\eta(\tilde\mu_t^{N,r_N}(\mathbb{R}))\partial_t g_t\mathrm{d}t\right|\underset{N\to\infty}{\longrightarrow}0.$$

From the constancy of $\tilde\mu_t^{N,r_N}(\mathbb{R})$ and an integration by parts, one has

$$-T_N^1+\int_0^\infty\int_{\mathbb{R}}\eta(\tilde\mu_t^{N,r_N}(\mathbb{R}))\partial_t g_t\mathrm{d}t=-\int_{\mathbb{R}}g_0\eta(H*\tilde\mu_0^{N,r_N}).$$

Another integration by parts yields

$$\begin{aligned}&2c_\alpha\sum_{i=1}^N\int_0^\infty w_t^i\chi_t^N\int_{\{|y|\le r_N\}}\varphi_t\left(X_t^{N,i,r_N}+\sigma_N y\right)-\varphi_t\left(X_t^{N,i,r_N}\right)-\sigma_N y\partial_x\varphi_t\left(X_t^{N,i,r_N}\right)\frac{\mathrm{d}y\mathrm{d}t}{|y|^{1+\alpha}}\\ =&-2c_\alpha\int_0^\infty\chi_t^N\int_{\{|y|\le r_N\}}\int_{\mathbb{R}}\left(g_t(x+\sigma_N y)-g_t(x)-\sigma_N y\partial_x g_t(x)\right)\eta(H*\tilde\mu_t^{N,r_N}(x))\frac{\mathrm{d}x\mathrm{d}y\mathrm{d}t}{|y|^{1+\alpha}}\\ &+2c_\alpha\int_0^\infty\chi_t^N\eta(\tilde\mu_t^{N,r_N}(\mathbb{R}))\int_{\{|y|\le r_N\}}\int_{\mathbb{R}}\left(g_t(x+\sigma_N y)-g_t(x)-\sigma_N y\partial_x g_t(x)\right)\frac{\mathrm{d}x\mathrm{d}y\mathrm{d}t}{|y|^{1+\alpha}}\\ =&-2c_\alpha\int_0^\infty\chi_t^N\int_{\{|y|\le r_N\}}\int_{\mathbb{R}}\left(g_t(x+\sigma_N y)-g_t(x)-\sigma_N y\partial_x g_t(x)\right)\eta(H*\tilde\mu_t^{N,r_N}(x))\frac{\mathrm{d}x\mathrm{d}y\mathrm{d}t}{|y|^{1+\alpha}}.\end{aligned}$$

Moreover, from the regularity of A and η, it holds that

$$w^i_{\tau_t}A'\left(H*\tilde\mu^{N,r_N}_{\tau_t}(X^{N,i,r_N}_{\tau_t})\right)=\psi\left(H*\tilde\mu^{N,r_N}_{\tau_t}(X^{N,i,r_N}_{\tau_t})\right)-\psi\left(H*\tilde\mu^{N,r_N}_{\tau_t}(X^{N,i,r_N}_{\tau_t}-)\right)+o\left(\frac{1}{N}\right),$$

so that

$$\mathbb{E}\left|2\sum_{i=1}^N\int_0^\infty w^i_{\tau_t}\chi_t^N\partial_x\varphi_t\left(X_t^{N,i,r_N}\right)A'\left(H*\tilde\mu^{N,r_N}_{\tau_t}(X^{N,i,r_N}_{\tau_t})\right)\mathrm{d}t+2\int_0^\infty\chi_t^N\int_{\mathbb{R}}\partial_x g_t\psi(H*\tilde\mu_t^{N,r_N})\mathrm{d}t\right|\underset{N\to\infty}{\to}0,$$

from an adaptation of Lemma 7.3.14 (replacing η by ψ in the definition of w_t^i). This concludes the proof of the first item of Lemma 7.3.9.

To prove the second item, observe that the change of variable $z=\sigma_N y$ yields, for $r_N\le\frac{1}{\sigma_N}$,

$$\left|2c_\alpha \int_0^\infty \chi_t^N \int_{\mathbb{R}} \int_{\{|y|\le r_N\}} \eta(H * \tilde{\mu}_t^{N,r_N}(x)) \left(g_t(x+\sigma_N y) - g_t(x) - \sigma_N y \partial_x g_t(x)\right) \frac{\mathrm{d}y\mathrm{d}x\mathrm{d}t}{|y|^{1+\alpha}}\right|$$
$$\le 2c_\alpha \sigma_N^\alpha \int_0^\infty \chi_t^N \int_{\mathbb{R}} \int_{\{|z|\le 1\}} \left|\eta(H * \tilde{\mu}_t^{N,r_N}(x)) \left(g_t(x+z) - g_t(x) - z\partial_x g_t(x)\right)\right| \frac{\mathrm{d}z\mathrm{d}x\mathrm{d}t}{|z|^{1+\alpha}}.$$

Proof (Proof of Lemma 7.3.10). First notice that

$$T_N^3 = \sum_{\kappa_i^N > 0} \int_0^\infty (1-\chi_t^N) \int_{\{|y|>r_N\}} \left(\int_{\mathbb{R}} \varphi_t \mathrm{d}\rho_{t-}^{y,i}\right) \nu^{i,r_N}(\mathrm{d}y, \mathrm{d}t),$$

with ρ defined by the following formula : ($\bar{\mu}_t^{y,i,N,r_N}$ being the measure obtained by moving in the expression of $\tilde{\mu}_t^{N,r_N}$ the particle X_t^{N,i,r_N} to the position $X_t^{N,i,r_N} + \sigma_N y$)

$$\rho_t^{y,i} = \partial_x \left(\eta(H * \bar{\mu}_t^{y,i,N,r_N}) - \eta(H * \tilde{\mu}_t^{N,r_N})\right).$$

To prove the second item in Lemma 7.3.10, we integrate by parts, and, using the definition of $\bar{\mu}^{y,i,N,r_N}$ and the compactness of the support of g, it holds that

$$\left|\int_{\mathbb{R}} \varphi_t \mathrm{d}\rho_t^{y,i}\right| = \left|\int_{\mathbb{R}} g_t \left(\eta(H * \bar{\mu}_t^{y,i,N,r_N}) - \eta(H * \tilde{\mu}_t^{N,r_N})\right)\right| \le K \frac{(\sigma_N y) \wedge 1}{N}, \tag{7.16}$$

so that

$$\mathbb{E}|T_N^3| \le K \int_0^\infty (1-\chi_t^N) \int_{\{|y|>r_N\}} (\sigma_N y) \wedge 1 \frac{\mathrm{d}y\mathrm{d}t}{|y|^{1+\alpha}} \le K(\sigma_N^\alpha + \sigma_N r_N^{1-\alpha}).$$

Now let us prove the first item of Lemma 7.3.10. Applying the same martingale argument as the one used to prove $\mathbb{E}|M_N| \to 0$, and using the upper bound K/N in (7.16), one has

$$\mathbb{E}\left|T_N^3 - 2c_\alpha \int_0^\infty (1-\chi_t^N) \int_{\{|y|>r_N\}} \left(\sum_{\kappa_i^N > t} \int_{\mathbb{R}} \varphi_t \mathrm{d}\rho_t^{y,i}\right) \frac{\mathrm{d}y\mathrm{d}t}{|y|^{1+\alpha}}\right|^2 \le \frac{K}{r_N^\alpha N}.$$

Let us give a more explicit expression for $\rho_t^{y,i}$. For simplicity, we denote

$$\tilde{w}_t^i = \mathbf{1}_{\kappa_i^N > t} \left[\eta\left(\frac{1}{N} \sum_{\substack{j\neq i \\ \kappa_j^N > t}} \gamma(X_0^j) \mathbf{1}_{X_t^{N,j,r_N} \le X_t^{N,i,r_N} + \sigma_N y} + \frac{\gamma(X_0^i)}{N}\right) - \eta\left(\frac{1}{N} \sum_{\substack{j\neq i \\ \kappa_j^N > t}} \gamma(X_0^j) \mathbf{1}_{X_t^{N,j,r_N} \le X_t^{N,i,r_N} + \sigma_N y}\right)\right]$$

and for $i \neq j$,

$$\tilde{w}_t^{i,j,\pm} = \mathbf{1}_{\kappa_i^N > t} \mathbf{1}_{\kappa_j^N > t} \left[\eta\left(\frac{1}{N} \sum_{\substack{k\neq j \\ \kappa_k^N > t}} \gamma(X_0^k) \mathbf{1}_{X_t^{N,k,r_N} \le X_t^{N,i,r_N}} \pm \frac{\gamma(X_0^j)}{N}\right) - \eta\left(\frac{1}{N} \sum_{\substack{k\neq j \\ \kappa_k^N > t}} \gamma(X_0^k) \mathbf{1}_{X_t^{N,k,r_N} \le X_t^{N,i,r_N}}\right)\right].$$

One can write

$$\begin{aligned}\rho_t^y := \sum_{\kappa_i^N>t} \rho_t^{y,i} = \sum_{\kappa_i^N>t} \tilde{w}_t^i \delta_{X_t^{N,i,r_N}+\sigma_N y} - \sum_{\kappa_i^N>t} w_t^i \delta_{X_t^{N,i,r_N}} \\ + \sum_{\kappa_i^N>t} \left(\sum_{\kappa_j^N>t} \left(\tilde{w}_t^{i,j,+} - w_t^i\right) \mathbf{1}_{X_t^{N,i,r_N}<X_t^{N,j,r_N}} \mathbf{1}_{X_t^{N,j,r_N}+\sigma_N y<X_t^{N,i,r_N}} \right) \delta_{X_t^{N,i,r_N}} \\ + \sum_{\kappa_i^N>t} \left(\sum_{\kappa_j^N>t} \left(\tilde{w}_t^{i,j,-} - w_t^i\right) \mathbf{1}_{X_t^{N,j,r_N}<X_t^{N,i,r_N}} \mathbf{1}_{X_t^{N,i,r_N}<X_t^{N,j,r_N}+\sigma_N y} \right) \delta_{X_t^{N,i,r_N}}.\end{aligned} \tag{7.17}$$

In this expression, the two first terms deal with particles jumping from the site X_t^{N,i,r_N} to the site $X_t^{N,i,r_N} + \sigma_N y$, while the third term corresponds to the jump from right to left of the particle labelled j above the particle labelled i and, conversely, the fourth term corresponds to the jumps of particle j from left to right over particle i. Notice that this last equality, as well as (7.18) below, only holds when each $X_t^{N,i,r_N} + \sigma_N y$ is distinct from all X_t^{N,j,r_N}. However, for all t, this condition holds dy-almost everywhere, which is enough for our purpose.

In the entropic formulation (7.7), the term that should appear for large jumps is given by

$$2c_\alpha \int_0^\infty \int_{\{|y|>r_N\}} \left(\int_{\mathbb{R}} \varphi_t \mathrm{d}\sigma_t^y \right) \frac{\mathrm{d}y\mathrm{d}t}{|y|^{1+\alpha}},$$

where

$$\begin{aligned}\sigma_t^y =& \partial_x \left(\eta'(H * \tilde{\mu}_t^{N,r_N}) \left(H * \tilde{\mu}_t^{N,r_N}(\cdot - \sigma_N y) - H * \tilde{\mu}_t^{N,r_N} \right) \right) \\ =& \frac{1}{N} \sum_{\kappa_i^N>t} \gamma(X_0^i) \eta' \left(H * \tilde{\mu}_t^{N,r_N}(X_t^{N,i,r_N} + \sigma_N y) \right) \delta_{X_t^{N,i,r_N}+\sigma_N y} - \frac{1}{N} \sum_{\kappa_i^N>t} \gamma(X_0^i) \eta' \left(H * \tilde{\mu}_t^{N,r_N}(X_t^{N,i,r_N}) \right) \delta_{X_t^{N,i,r_N}} \\ &+ \sum_{\kappa_i^N>t} \left(H * \tilde{\mu}_t^{N,r_N}(X_t^{N,i,r_N} - \sigma_N y) - H * \tilde{\mu}_t^{N,r_N}(X_t^{N,i,r_N}) \right) \\ &\qquad \times \left(\eta' \left(H * \tilde{\mu}_t^{N,r_N}(X_t^{N,i,r_N}) \right) - \eta' \left(H * \tilde{\mu}_t^{N,r_N}(X_t^{N,i,r_N} -) \right) \right) \delta_{X_t^{N,i,r_N}}.\end{aligned} \tag{7.18}$$

When computing the difference $\rho_t^y - \sigma_t^y$ integrated against some bounded function, using Taylor expansions for η, one can check that, up to an error term of order $\mathcal{O}\left(\frac{1}{N}\right)$ the first terms in the right hand side of (7.17) and (7.18) cancel each other, the second terms as well, and so does the sum of the two last term in (7.17) with the last one in (7.18). Consequently,

$$\left| \int_0^\infty (1-\chi_t^N) \int_{\{|y|>r_N\}} \left(\int_{\mathbb{R}} \varphi_t \mathrm{d}\rho_t^y \right) \frac{\mathrm{d}y\mathrm{d}t}{|y|^{1+\alpha}} - \int_0^\infty (1-\chi_t^N) \int_{\{|y|>r_N\}} \left(\int_{\mathbb{R}} \varphi_t \mathrm{d}\sigma_t^y \right) \frac{\mathrm{d}y\mathrm{d}t}{|y|^{1+\alpha}} \right| \leq \frac{K}{N r_N^\alpha}.$$

This concludes the proof.

Proof (Proof of Lemma 7.3.11). For a time ζ_m of the form kh_N, no particle moved in the interval (ζ_{m-1}, ζ_m), so that $w_{\zeta_m}^i - w_{\zeta_{m-1}}^i = 0$, unless the particle labelled i has been killed at time ζ_m. Hence,

$$\begin{aligned}T_N^4 =& \sum_{i=1}^N \sum_{\substack{\zeta_m \text{ of the} \\ \text{form } kh_N}} (w_{\zeta_m}^i - w_{\zeta_{m-1}}^i) \varphi_{\zeta_m}(X_{\zeta_m}^{N,i,r_N}) \\ =& - \sum_{\substack{\zeta_m \text{ of the} \\ \text{form } kh_N}} \sum_{\kappa_i^N=\zeta_m} w_{\zeta_{m-1}}^i \varphi_{\zeta_m}(X_{\zeta_m}^{N,i,r_N}).\end{aligned}$$

This sum is actually a sum over pairs of close particles with opposite signs, thus

$$\begin{aligned}|T_N^4| &= \left| \sum_{\substack{\zeta_m \text{ of the} \\ \text{form } kh_N}} \sum_{\substack{\text{pairs } \{i,j\} \text{ of particles} \\ \text{killed at time } \zeta_m}} \left(w^i_{\zeta_{m-1}} \varphi_{\zeta_m}(X^{N,i,r_N}_{\zeta_m}) + w^j_{\zeta_{m-1}} \varphi_{\zeta_m}(X^{N,j,r_N}_{\zeta_m}) \right) \right| \\ &\leq \sum_{\substack{\zeta_m \text{ of the} \\ \text{form } kh_N}} \sum_{\substack{\text{pairs } \{i,j\} \text{ of particles} \\ \text{killed at time } \zeta_m}} \left| w^i_{\zeta_{m-1}} + w^j_{\zeta_{m-1}} \right| \|\varphi\|_\infty + \left| w^j_{\zeta_{m-1}} \right| \left| \varphi_{\zeta_m}(X^{N,i,r_N}_{\zeta_m}) - \varphi_{\zeta_m}(X^{N,j,r_N}_{\zeta_m}) \right| \\ &\leq K \left(\frac{1}{N} + \varepsilon_N \right).\end{aligned}$$

Indeed, a couple (i,j) of killed particles is such that $|X^{N,i,r_N}_{\zeta_m} - X^{N,j,r_N}_{\zeta_m}| \leq \varepsilon_N$ and is made of particles with opposite signs, so that

$$|w^i_{\zeta_{m-1}} + w^j_{\zeta_{m-1}}| = \left| \left(\gamma(X^i_0) + \gamma(X^j_0) \right) \eta'(H * \tilde{\mu}^{N,r_N}_{\zeta_{m-1}}(X^{N,i,r_N}_{\zeta_{m-1}})) + \mathcal{O}\left(\frac{1}{N^2} \right) \right| \leq \frac{K}{N^2}.$$

Proof (Proof of Lemma 7.3.12). Notice that from independence of the increments, denoting by $\mathcal{L}^{\leq r}$ a Lévy process with Lévy measure $c_\alpha \mathbf{1}_{|y| \leq r} \frac{\mathrm{d}y}{|y|^{1+\alpha}}$, one has

$$\begin{aligned}\mathbb{P}(B_N) &= \mathbb{P}(\sigma_N |\mathcal{L}^{\leq r_N}_{h_N}| \leq \varepsilon_N/4)^{NT/h_N} \\ &= \left(1 - \mathbb{P}\left(\sigma_N r_N |\mathcal{L}^{\leq 1}_{h_N r_N^{-\alpha}}| \geq \varepsilon_N/4 \right) \right)^{NT/h_N}.\end{aligned}$$

Since the Lévy measure $c_\alpha \mathbf{1}_{|y| \leq 1} \frac{\mathrm{d}y}{|y|^{1+\alpha}}$ has compact support, the random variables $\mathcal{L}^{\leq 1}_t$ have exponential moments, and Chernov's inequality yields

$$\mathbb{P}\left(\sigma_N r_N |\mathcal{L}^{\leq 1}_{h_N r_N^{-\alpha}}| \geq \varepsilon_N/4 \right) \leq \mathbb{E}\left(\exp\left(\left| \mathcal{L}^{\leq 1}_{h_N r_N^{-\alpha}} \right| \right) \right) \exp\left(-\varepsilon_N/4\sigma_N r_N \right) = \exp\left(K h_N r_N^{-\alpha} - \varepsilon_N/4\sigma_N r_N \right),$$

where the constant K does not depend on N.

In the Brownian case $\alpha = 2$, we use the tail estimate $\int_M^\infty e^{-x^2} \mathrm{d}x \leq K e^{-M^2}$ for positive M.

Lemma 7.3.13 *Let $a_1 \leq \ldots \leq a_N$ and $b_1 \leq \ldots \leq b_N$ be two nondecreasing sequences of reals numbers. Then the quantity $\sum_{i=1}^N a_i b_{\sigma(i)}$ for some permutation σ is maximal when $\sigma(i) = i$ for all i.*

Proof. From optimal transportation theory (see [71, page 75]), the quantity $\sum_{i=1}^N (a_i - b_{\sigma(i)})^2$ is minimal when σ is the identity. Expanding the square, we see that $\sum_{i=1}^N (a_i - b_{\sigma(i)})^2 = \sum_{i=1}^N (a_i^2 + b_i^2) - 2\sum_{i=1}^N a_i b_{\sigma(i)}$. Thus, $\sum_{i=1}^N a_i b_{\sigma(i)}$ is maximal if and only if $\sum_{i=1}^N (a_i - b_{\sigma(i)})^2$ is minimal, concluding the proof.

Lemma 7.3.14 *Let f be some bounded function with compact support on $[0,\infty) \times \mathbb{R}$ which is smooth with respect to the space variable. If h_N goes to zero and σ_N is bounded, then*

$$\lim_{N \to \infty} \mathbb{E}\left| \sum_{\kappa_i^N > 0} \int_0^\infty \left(w^i_t - w^i_{\tau_t} \right) f_t\left(X^{N,i,r_N}_t \right) \mathrm{d}t \right| = 0.$$

Proof. First notice that when t is not in an interval $[kh_N, (k+1/2)h_N]$, one has $w^i_t = w^i_{\tau_t}$, since no particle moved between τ_t and t. Then, one can write, from the assumptions on f,

$$\left|\sum_{\kappa_i^N>0}\int_0^\infty (w_t^i - w_{\tau_t}^i)\, f_t\left(X_t^{N,i,r_N}\right)\mathrm{d}t\right| \le \left|\sum_{\kappa_i^N>0}\int_0^T \chi_t^N\left(w_t^i f_t(X_t^{N,i,r_N}) - w_{\tau_t}^i f_t(X_{\tau_t}^{N,i,r_N})\right)\mathrm{d}t\right| + \frac{K}{N}\sum_{\kappa_i^N>0}\int_0^T \chi_t^N\left|X_t^{N,i,r_N} - X_{\tau_t}^{N,i,r_N}\right| \wedge 1\mathrm{d}t.$$

Integrating by parts, one can see that

$$\begin{aligned}
&\left|\sum_{\kappa_i^N>0}\int_0^T \chi_t^N\left(w_t^i f_t(X_t^{N,i,r_N}) - w_{\tau_t}^i f_t(X_{\tau_t}^{N,i,r_N})\right)\mathrm{d}t\right| \\
=&\left|\int_0^T \chi_t^N \int_{\mathbb{R}}\left(\eta\left(H * \tilde{\mu}_t^{N,r_N}(x)\right) - \eta\left(H * \tilde{\mu}_{\tau_t}^{N,r_N}(x)\right)\right)\partial_x f_t(x)\mathrm{d}x\mathrm{d}t\right| \\
\le&\frac{K}{N}\int_0^T \chi_t^N \int_{\mathbb{R}}\left(\sum_{\kappa_i^N>0}\mathbf{1}_{X_t^{N,i,r_N}\le x< X_{\tau_t}^{N,i,r_N}} + \mathbf{1}_{X_{\tau_t}^{N,i,r_N}\le x< X_t^{N,i,r_N}}\right)\partial_x f_t(x)\mathrm{d}x\mathrm{d}t \\
\le&\frac{K}{N}\int_0^T \chi_t^N \sum_{\kappa_i^N>0}\left|X_t^{N,i,r_N} - X_{\tau_t}^{N,i,r_N}\right| \wedge 1\mathrm{d}t
\end{aligned}$$

We conclude the proof by writing

$$\begin{aligned}
\mathbb{E}\frac{1}{N}\int_0^T \chi_t^N \sum_{\kappa_i^N>0}\left|X_t^{N,i,r_N} - X_{\tau_t}^{N,i,r_N}\right| \wedge 1\mathrm{d}t &= \mathbb{E}\int_0^T \chi_t^N \mathbf{1}_{\kappa_1^N>0}\left|X_t^{N,1,r_N} - X_{\tau_t}^{N,1,r_N}\right| \wedge 1\mathrm{d}t \\
&\le T\left(h_N \sup_{[-1,1]}|A'| + \mathbb{E}\left((\sigma_N|\Lambda_{h_N}^{N,1,r_N}|)\wedge 1\right)\right).
\end{aligned}$$

This last quantity tends to zero when h_N goes to 0.

7.4 Numerical results

In this section, we illustrate our convergence results by some numerical simulations. We simulated the solution to the fractional and the inviscid Burgers equations

$$\partial_t u + \frac{1}{2}\partial_x(u^2) + \sigma^\alpha(-\Delta)^{\frac{\alpha}{2}} = 0 \quad \text{and} \quad \partial_t u + \frac{1}{2}\partial_x(u^2) = 0,$$

corresponding to the choice $A(x) = x^2/2$, with different values for the parameter α.

One can find an explicit exact solution to the inviscid Burgers equation (see [46]) and we compare the result of the simulation to this exact solution in the vanishing viscosity setting. However, to our knowledge, no explicit solutions exist in the case of a positive viscosity coefficient for $\alpha < 2$, so that we have to compare the result of our simulation with the one given by another numerical method. Here, we use a deterministic method, introduced by Droniou in [29].

7.4.1 Constant viscosity ($\sigma_N = \sigma$)

We give three examples of approximation to the viscous conservation law. On Figures 7.1, 7.2 and 7.3, we show the approximation of the solution starting at $\mathbf{1}_{[-0.2,0.2]}$ to the viscous conservation law with respective index $\alpha = 1.5$, $\alpha = 1$ and $\alpha = 0.1$ and diffusion coefficient $\sigma = 1$ using $N = 1000$ particles, with parameters $h = 0.01$ and $\varepsilon = 0.04$ at simulation times 0.25, 0.5, 0.75

and 1. The continuous line is the simulated solution, and the dotted line is the "exact" solution obtained with the determistic scheme of [29] using small time and space steps.

We now investigate the convergence rate of the error, that is the Riemann sum on the discretization grid associated to the integral in Theorems 7.3.1, 7.3.2 and 7.3.3. On Figure 7.4 is depicted the logarithmic plot of the error as a function of N where we used the relation $h_N = 10/N$, and $\varepsilon_N = 40/N$, with N ranging from 10 to 10000, in the three cases $\alpha = 0.5$, 1 and 1.5. In the case $\alpha < 1$, this relation between N, h_N and ε_N satisfies the condition of Theorem 7.3.1. These pictures make us expect a convergence rate of $\frac{1}{\sqrt{N}}$, corresponding to the optimal rate analyzed theoretically in [17,18], in the case $\alpha = 2$, without killing.

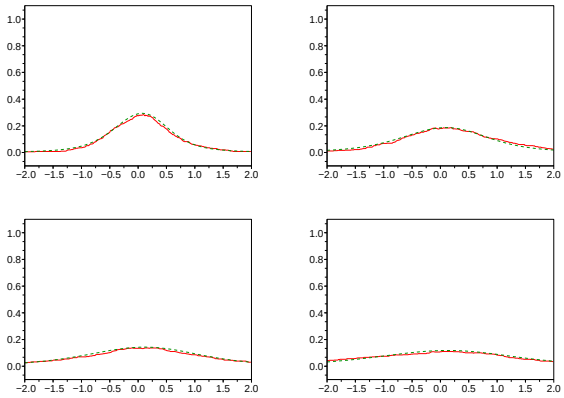

Fig. 7.1. Approximation of the conservation law with index $\alpha = 1.5$.

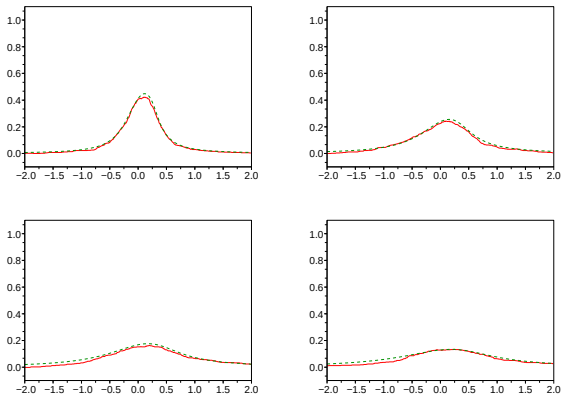

Fig. 7.2. Approximation of the conservation law with index $\alpha = 1$.

Behaviour as $h \to 0$

We give in Figure 7.5 the approximation error in terms of the time step h, for a fixed number of particles, in logarithmic plot. We set the parameter ε to be equal to $4h$ so that the condition of Theorem 7.3.1 is satified. We took $N = 340000$ and $\sigma = 1$. We set $\alpha = 0.5$, $\alpha = 1$ and $\alpha = 1.5$ respectively. The different parameters h range from 1 to 2^{-8}. In [17,18] it is shown, in case $\alpha = 2$

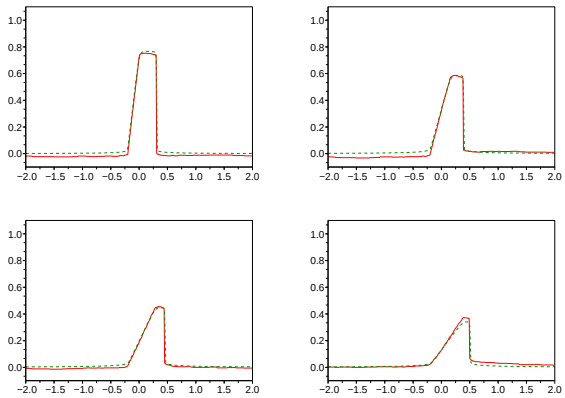

Fig. 7.3. Approximation of the conservation law with index $\alpha = 0.1$.

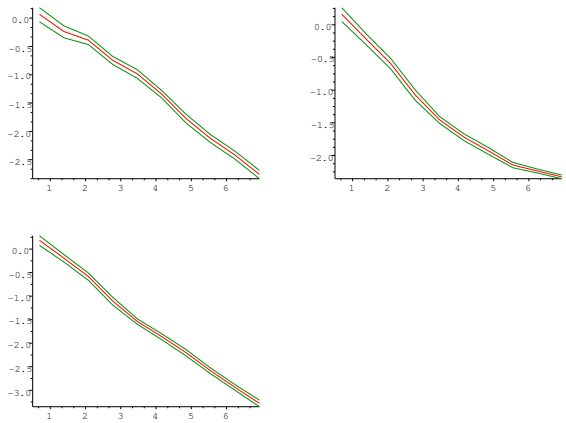

Fig. 7.4. Error in the approximation of the conservation law with index $\alpha = 0.5$, 1 and 1.5 as N tends to infinity, in log-log plot. The respective slopes are -0.46, -0.41 and -0.56. The upper and lower lines show the 95% confidence interval.

and the initial condition is monotonic, that the error is of order h. In view of Figure 7.5, it seems that the convergence rate is still of order h, even for $\alpha < 2$ and any initial condition with bounded variation.

7.4.2 Vanishing viscosity ($\sigma_N \to 0$)

We consider the Burgers equation

$$\partial_t v = \partial_x (u^2/2)$$

with initial condition $u_0(x) = \mathbf{1}_{[-3,-2]} - \mathbf{1}_{[2,3]}$, which is the cumulative distribution function of the measure $\delta_{-3} - \delta_{-2} + \delta_2 - \delta_3$. In that case, the solution of the Burgers equation is explicit and given by the expression

$$u(t,x) = \min\left(\frac{x+3}{t}, 1\right) \mathbf{1}_{[-3,\min(-2+\frac{t}{2}, -3+\sqrt{2t}, 0)]} + \max\left(\frac{x-3}{t}, -1\right) \mathbf{1}_{[\max(2-\frac{t}{2}, 3-\sqrt{2t}, 0), 3]}.$$

We compare the function u to the function obtained by running the Euler scheme with a small diffusion coefficient σ. One can expect the approximation to be better for large values of α. Indeed,

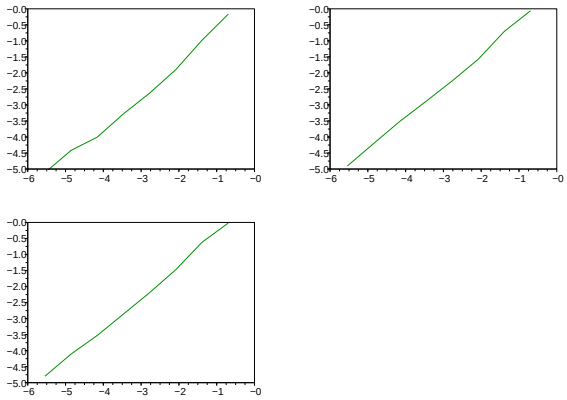

Fig. 7.5. Log-log plot of the error as h tends to zero, with a fixed number of particles, at respectively $\alpha = 0.5$, 1 and 1.5. The slopes are equal to 1 up to an error of 0.01.

for small values of α, the particles tend to jump very far away, and subsequently "disappear" from the simulation. The consequence of this behaviour is that the solution is somehow decreased by a multiplicative constant.

For large values of α, the approximation is quite good, even for not so small diffusion coefficients. Figure 7.6 gives the result of the simulation of the Euler scheme with parameters $\alpha = 1.5$, $\varepsilon = 0.04$, $\sigma = 0.1$ and $h = 0.01$, at the different times 2, 4, 6 and 8 for $N = 10000$ particles. Figure 7.7 gives the same simulation for $\alpha = 1$. In the case $\alpha < 1$, and especially when α is small, one need to take a very small value for the diffusion coefficient in order to have a reasonable approximation of the solution. Indeed, the approximation depicted on the Figure 7.8 is the appproximation of the solution at times 2, 4, 6 and 8 for diffusion coefficient $\sigma = 10^{-4}$. Here, we used 10000 particles killed at a distance $\varepsilon = 0.01$, the time step being $h = 0.01$. On Figure 7.9 we show the same simulation, with diffusion coefficient changed to $\sigma = 10^{-12}$.

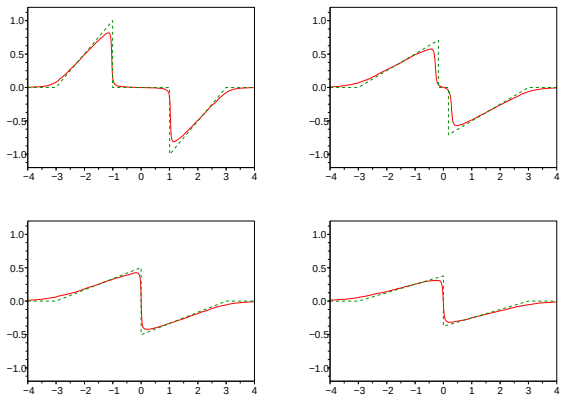

Fig. 7.6. Approximation of the inviscid conservation law by a fractional Euler scheme with index $\alpha = 1.5$ and diffusion coefficient 0.1.

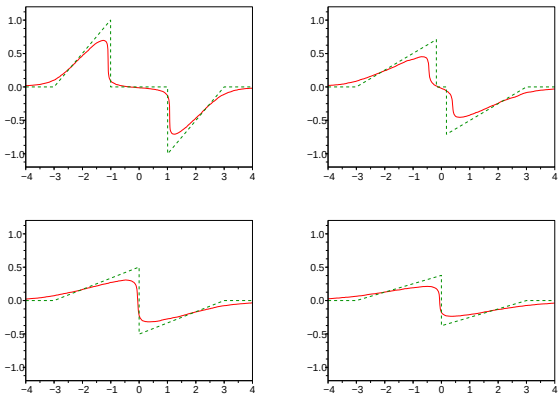

Fig. 7.7. Approximation of the inviscid conservation law by a fractional Euler scheme with index $\alpha = 1$ and diffusion coefficient 0.1.

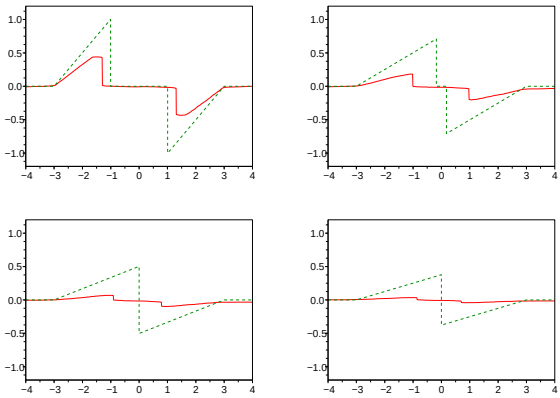

Fig. 7.8. Approximation of the inviscid conservation law by a fractional Euler scheme with index $\alpha = 0.1$ and diffusion coefficient 10^{-4}.

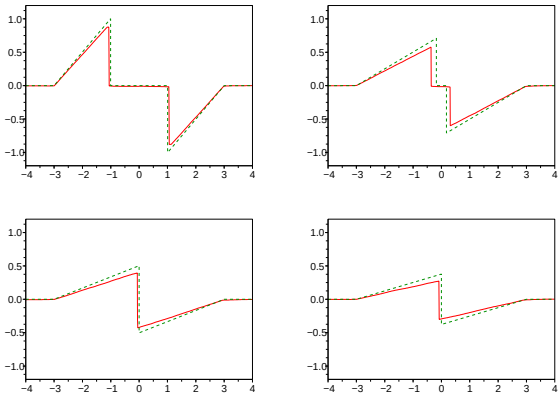

Fig. 7.9. Approximation of the inviscid conservation law by a fractional Euler scheme with index $\alpha = 0.1$ and diffusion coefficient 10^{-12}.

Partie V

Bibliographie

Bibliographie

[1] R. Adams, *Sobolev spaces* (Academic Press, 1978).

[2] D. Aldous, Stopping times and tightness, *Ann. Probability* **6**(2) (1978) 335–340.

[3] N. Alibaud, Entropy formulation for fractal conservation laws, *J. Evol. Equ.* **7**(1) (2007) 145–175.

[4] N. Alibaud and B. Andreianov, Non-uniqueness of weak solutions for the fractal Burgers equation, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **27**(4) (2010) 997–1016.

[5] N. Alibaud, J. Droniou, and J. Vovelle, Occurrence and non-appearance of shocks in fractal Burgers equations, *J. Hyperbolic Differ. Equ.* **4**(3) (2007) 479–499.

[6] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer, *Sur les inégalités de Sobolev logarithmiques* (Société Mathématique de France, 2000). In French.

[7] R. Assaraf and M. Caffarel, Zero-variance zero-bias principle for observables in quantum Monte Carlo : Application to forces, *J. Chem. Phys.* **119**(20) (2003).

[8] D. Bakry, L'hypercontractivité et son utilisation en théorie des semigroupes, In *Lectures on probability theory (Saint-Flour, 1992)*, volume 1581 of *Lecture Notes in Math.* (Springer, Berlin, 1994), pp. 1–114.

[9] V. Bally and D. Talay, The Euler scheme for stochastic differential equations : error analysis with Malliavin calculus, *Math. Comput. Simulation* **38**(1-3) (1995) 35–41. Probabilités numériques (Paris, 1992).

[10] V. Bally and D. Talay, The law of the Euler scheme for stochastic differential equations. I. Convergence rate of the distribution function, *Probab. Theory Related Fields* **104**(1) (1996) 43–60.

[11] V. Bally and D. Talay, The law of the Euler scheme for stochastic differential equations. II. Convergence rate of the density, *Monte Carlo Methods Appl.* **2**(2) (1996) 93–128.

[12] P. Biler, T. Funaki, and W. A. Woyczynski, Fractal Burgers equations, *J. Differential Equations* **148**(1) (1998) 9–46.

[13] P. Biler, G. Karch, and W. A. Woyczyński, Asymptotics for conservation laws involving Lévy diffusion generators, *Studia Math.* **148**(2) (2001) 171–192.

[14] M. Bossy, Optimal rate of convergence of a stochastic particle method to solutions of 1D viscous scalar conservation laws, *Math. Comp.* **73**(246) (2004) 777–812 (electronic).

[15] M. Bossy, L. Fezoui, and S. Piperno, Comparison of a stochastic particle method and a finite volume deterministic method applied to Burgers equation, *Monte Carlo Methods Appl.* **3**(2) (1997) 113–140.

[16] M. Bossy, J.F. Jabir, and D. Talay, On conditional McKean Lagrangian stochastic models, *Probability Theory and Related Fields* (2010) 1–33.

[17] M. Bossy and D. Talay, Convergence rate for the approximation of the limit law of weakly interacting particles : application to the Burgers equation, *Ann. Appl. Probab.* **6**(3) (1996) 818–861.

[18] M. Bossy and D. Talay, A stochastic particle method for the McKean-Vlasov and the Burgers equation, *Math. Comp.* **66**(217) (1997) 157–192.

[19] Y. Brenier and E. Grenier, Sticky particles and scalar conservation law, *SIAM J. Numer. Anal.* **35**(6) (1998) 2317–2328 (electronic).

[20] H. Brézis, *Analyse fonctionnelle. Théorie et applications*, Collection Mathématiques appliquées pour la maîtrise (Masson, Paris, 1983). In French.

[21] E. Cancès, B. Joudain, and T. Lelièvre, Quantum Monte Carlo simulations of fermions. A mathematical analysis of the fixed-node approximation, *Math. Models Methods Appl. Sci.* **16**(9) (2006) 1403–1440.

[22] C.-H. Chan and M. Czubak, Regularity of solutions for the critical N-dimensional Burgers' equation, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **27**(2) (2010) 471–501.

[23] C.-H. Chan, M. Czubak, and L. Silvestre, Eventual regularization of the slightly supercritical fractional Burgers equation, *Discrete Contin. Dyn. Syst.* **27**(2) (2010) 847–861.

[24] C. Chipot and A. Pohorille (Eds.), *Free Energy Calculations*, volume 86 of *Springer Series in Chemical Physics* (Springer, 2007).

[25] E. Darve and A. Pohorille, Calculating free energy using average forces, *J. Chem. Phys.* **115** (2001) 9169–9183.

[26] R. Dautray and P.L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology* (Springer Verlag, 1999).

[27] P. Del Moral, *Feynman-Kac Formulae : Genealogical and Interacting Particle Systems with Applications* (Springer-Verlag, 2004).

[28] A. Dermoune, Propagation and conditional propagation of chaos for pressureless gas equations, *Probab. Theor. Relat. Fields* **126** (2003) 459–479.

[29] J. Droniou, A numerical method for fractal conservation law, *Math. Comp.* **79** (2010) 95–124.

[30] J. Droniou, T. Gallouet, and J. Vovelle, Global solution and smoothing effect for a non-local regularization of a hyperbolic equation, *J. Evol. Equ.* **3**(3) (2003) 499–521. Dedicated to Philippe Bénilan.

[31] J. Droniou and C. Imbert, Fractal first-order partial differential equations, *Arch. Ration. Mech. Anal.* **182**(2) (2006) 299–331.

[32] J. Hénin and C. Chipot, Overcoming free energy barriers using unconstrained molecular dynamics simulations, *J. Chem. Phys.* **121** (2004) 2904–2914.

[33] N. Jacob, A. Potrykus, and J.-L. Wu, Solving a nonlinear pseudo-differential equation of Burgers type, *Stoch. Dyn.* **8**(4) (2008) 613–624.

[34] B. Jourdain, Diffusion processes associated with nonlinear evolution equations for signed measures, *Methodol. Comput. Appl. Probab.* **2**(1) (2000) 69–91.

[35] B. Jourdain, Particules collantes signées et lois de conservation scalaires 1D, *C. R. Math. Acad. Sci. Paris* **334**(3) (2002) 233–238.

[36] B. Jourdain, Probabilistic characteristics method for a one-dimensional inviscid scalar conservation law., *Ann. Appl. Probab.* **12**(1) (2002) 334–360.

[37] B. Jourdain, T. Lelièvre, and R. Roux, Existence, uniqueness and convergence of a particle approximation for the adaptive biasing force process, *ESAIM : M2AN* **44**(5) (2010) 831–865.

[38] B. Jourdain and F. Malrieu, Propagation of chaos and Poincaré inequalities for a system of particles interacting through their CDF, *Ann. Appl. Probab.* **18**(5) (2008) 1706–1736.

[39] B. Jourdain, S. Méléard, and W. A. Woyczyński, Probabilistic approximation and inviscid limits for one-dimensional fractional conservation laws, *Bernoulli* **11**(4) (2005) 689–714.

[40] I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*, second ed., volume 113 of *Graduate Texts in Mathematics* (Springer-Verlag, New York, 1991).

[41] W. Kliemann, Recurrence and invariant measures for degenerate diffusions, *Ann. Probab.* **15**(2) (1987) 690–707.

[42] V. N. Kolokoltsov, Nonlinear Markov semigroups and interacting Lévy type processes, *J. Stat. Phys.* **126**(3) (2007) 585–642.

[43] V. N. Kolokoltsov, *Nonlinear Markov processes and kinetic equations*, volume 182 of *Cambridge Tracts in Mathematics* (Cambridge University Press, Cambridge, 2010).

[44] S.N. Kruzhkov, First order quasilinear equations in several independant variables, *Mat. Sb.* **81** (1970) 228–255.

[45] N.V. Krylov and M. Röckner, Strong solutions of stochastic equations with singular time dependent drift, *Probab. Theor. Relat. Fields* **131** (2005) 154–196.

[46] P. D. Lax, Hyperbolic systems of conservation laws. II, *Comm. Pure Appl. Math.* **10** (1957) 537–566.

[47] T. Lelièvre, M. Rousset, and G. Stoltz, Computation of free energy profiles with parallel adaptive dynamics, *J. Chem. Phys* **126** (2007) 134111.

[48] T. Lelièvre, M. Rousset, and G. Stoltz, Long-time convergence of an adaptive biasing force method, *Nonlinearity* **21** (2008) 1155–1181.

[49] T. Lelièvre, M. Rousset, and G. Stoltz, *Free energy computation : a mathematical perspective* (Imperial College Press, 2010).

[50] J.L. Lions, *Quelques méthodes de résolution des problèmes aux limites non-linéaires* (Dunod, 1969). In French.

[51] J.L. Lions and E. Magenes, *Problèmes aux limites non homogènes et applications* (Dunod, Paris, 1968-1970). In French.

[52] G. Maruyama and H. Tanaka, Ergodic property of N-dimensional recurrent Markov processes, *Memoirs of the Faculty of Science, Kyūshū University. Series A, Mathematics* **13**(2) (1959) 157–172.

[53] H. P. McKean, Jr., Propagation of chaos for a class of non-linear parabolic equations., In *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)* (Air Force Office Sci. Res., Arlington, Va., 1967), pp. 41–57.

[54] S. Méléard, Asymptotic behaviour of some interacting particle systems ; McKean-Vlasov and Boltzmann models, In *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995)*, volume 1627 of *Lecture Notes in Math.* (Springer, Berlin, 1996), pp. 42–95.

[55] P. Metzner, Ch. Schütte, and E. Vanden-Eijnden, Illustration of transition path theory on a collection of simple examples, *J. Chem. Phys* **125** (2006) 084110.

[56] S. Pal and J. Pitman, One-dimensional Brownian particle systems with rank-dependent drifts, *Ann. Appl. Probab.* **18**(6) (2008) 2179–2207.

[57] E. Priola. Pathwise uniqueness for singular SDEs driven by stable processes, 2010. Available at `http://arxiv.org/abs/1005.4237`.

[58] P. Protter and D. Talay, The Euler scheme for Lévy driven stochastic differential equations, *Ann. Probab.* **25**(1) (1997) 393–423.

[59] M. Rousset, On the control of an interacting particle approximation of Schrödinger ground states, *SIAM J. Math. Anal.* **38**(3) (2006) 824–844.

[60] M. Rousset and G. Stoltz, Equilibrium sampling from nonequilibrium dynamics, *J. Stat. Phys.* **123**(6) (2006) 1251–1272.

[61] G. Royer, *Une initiation aux inégalités de Sobolev logarithmiques*, volume 5 of *Cours Spécialisés [Specialized Courses]* (Société Mathématique de France, Paris, 1999).

[62] D. Serre, *Systèmes de lois de conservation. I*, Fondations. [Foundations] (Diderot Editeur, Paris, 1996). Hyperbolicité, entropies, ondes de choc. [Hyperbolicity, entropies, shock waves].

[63] D. Stanescu, D. Kim, and W. A. Woyczynski, Numerical study of interacting particles approximation for integro-differential equations, *J. Comput. Phys.* **206**(2) (2005) 706–726.

[64] A.S. Sznitman, Topics in propagation of chaos, *Lecture notes in mathematics* **1464** (1989).

[65] D. Talay and L. Tubaro, Expansion of the global error for numerical schemes solving stochastic differential equations, *Stochastic Anal. Appl.* **8**(4) (1990) 483–509 (1991).

[66] D. Talay and O. Vaillant, A stochastic particle method with random weights for the computation of statistical solutions of McKean-Vlasov equations, *Anal.Appl.Probab.* **13**(1) (2003) 140–180.

[67] R. Temam, *Navier-Stokes equations and nonlinear functionnal analysis* (North Holland, Amsterdam, 1979).

[68] J. Toulouse, R. Assaraf, and C.J. Umrigar, Zero-variance zero-bias quantum monte carlo estimators of the spherically and system-averaged pair density, *J. Chem. Phys.* **126** (2007) 244112.

[69] V.C. Tran, A wavelet particle approximation for McKean-Vlasov and 2D-Navier-Stokes statistical solutions, *Stochastic Processes and their Applications* **118** (2008) 284–318.

[70] A.B. Tsybakov, *Introduction à l'estimation non-paramétrique* (Springer, 2004). In French.

[71] C. Villani, *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics* (American Mathematical Society, 2003).

[72] W. A. Woyczyński, Lévy processes in the physical sciences, In *Lévy processes* (Birkhäuser Boston, Boston, MA, 2001), pp. 241–266.

[73] X.-C. Zhang. Discontinuous stochastic differential equations driven by Lévy processes, 2010. Available at `http://arxiv.org/abs/1011.5600`.